Lumped finite elements for reaction–cross-diffusion systems on stationary surfaces

ARTICLE IN PRESS

Computers and Mathematics with Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Lumped finite elements for reaction–cross-diffusion systems on stationary surfaces[☆]

Massimo Frittelli [a,*], Anotida Madzvamuse [b], Ivonne Sgura [a], Chandrasekhar Venkataraman [c]

[a] Dipartimento di Matematica e Fisica E. De Giorgi, Università del Salento, via per Arnesano, I-73100 Lecce, Italy

[b] University of Sussex, School of Mathematical and Physical Sciences, Department of Mathematics, Pevensey III, 5C15, Brighton, BN1 9QH, United Kingdom

[c] School of Mathematics and Statistics, University of St Andrews, Fife, KY16 9SS, United Kingdom

ARTICLE INFO

Article history:
Received 28 October 2016
Received in revised form 24 July 2017
Accepted 26 July 2017
Available online xxxx

Keywords:
Surface finite elements
Mass lumping
Invariant region
Reaction–cross-diffusion
Convergence analysis
Pattern formation

ABSTRACT

We consider a lumped surface finite element method (LSFEM) for the spatial approximation of reaction–diffusion equations on closed compact surfaces in $\mathbb{R}^3$ in the presence of cross-diffusion. We provide a fully-discrete scheme by applying the Implicit–Explicit (IMEX) Euler method. We provide sufficient conditions for the existence of polytopal invariant regions for the numerical solution after spatial and full discretisations. Furthermore, we prove optimal error bounds for the semi- and fully-discrete methods, that is the convergence rates are quadratic in the meshsize and linear in the timestep. To support our theoretical findings, we provide two numerical tests. The first test confirms that in the absence of lumping numerical solutions violate the invariant region leading to blow-up due to the nature of the kinetics. The second experiment is an example of Turing pattern formation in the presence of cross-diffusion on the sphere.

1. Introduction

In recent years, there has been a remarkable increase in the theoretical analysis of mathematical models of reaction–diffusion type in the presence of *cross-diffusion*. Cross-diffusion is a process in which the gradient in the concentration or density of one chemical or biological species induces a flux, either linearly or nonlinearly, of another species. In molecular biology, cross-diffusion processes appear in multicomponent systems containing at least two solute components [1,2]. Multicomponent systems containing nanoparticles, surfactants, polymers and other macromolecules in solution play an important role in industrial applications and biological functions [1]. In developmental biology, recent experimental findings demonstrate that cross-diffusion can be quite significant in generating spatial structure [3]. The effects of cross-diffusion on models for pattern formation have been studied in many theoretical papers, such as [4]. Apart from pattern formation in developmental biology, other applications of reaction–cross-diffusion systems include cancer motility [5], finance [6] and biofilms [7]. The introduction of cross-diffusion in standard reaction–diffusion models has been shown to prevent blow-up phenomena that are associated with reaction–diffusion systems in the absence of cross-diffusion [8]. It must be noted that

☆ This article belongs to SDS2016.

* Corresponding author.
E-mail addresses: massimo.frittelli@unisalento.it (M. Frittelli), A.Madzvamuse@sussex.ac.uk (A. Madzvamuse), ivonne.sgura@unisalento.it (I. Sgura), cv28@st-andrews.ac.uk (C. Venkataraman).

http://dx.doi.org/10.1016/j.camwa.2017.07.044

the concept of cross-diffusion includes well-known processes such as chemo- and haptotaxis [5]. In this paper we consider reaction–cross-diffusion systems (RCDSs) of $r \geq 1$ equations on a stationary surface of the form

$$\begin{cases} \dfrac{\partial u_m}{\partial t} - \displaystyle\sum_{k=1}^{r} d_{mk}\Delta_\Gamma u_k = f_m(u_1,\ldots,u_r), & \text{in } \Gamma \times (0,T), \\ u_m(\boldsymbol{x},0) = u_{0,m}(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \Gamma, & m = 1,\ldots,r, \end{cases} \tag{1}$$

where Γ is a smooth stationary orientable surface of co-dimension one in $\mathbb{R}^3$ without boundary, Δ_Γ is the Laplace–Beltrami operator on Γ (which is defined as the tangential divergence of the tangential gradient, see [9] for definitions), d_{ij} are any real diffusion and cross-diffusion coefficients such that the diffusion matrix $\boldsymbol{D} = (d_{ij})$ is positive definite, $f_1,\ldots,f_r$ are $\mathcal{C}^2(\Sigma;\mathbb{R})$ reaction kinetics on an compact set $\Sigma \subset \mathbb{R}^r$ and an initial condition $(u_{0,1},\ldots,u_{0,r}) \in \mathcal{C}^2(\Gamma)$ is given.

For many RCDSs, an important property is the existence of invariant regions. From a modelling point of view, it is useful to know that a given model possesses an invariant region. For real applications, solutions for RCDSs are usually meaningful as long as they range within a limited set of values and an invariant region could provide an a-priori bound on the analytical solution which can be helpful, for instance, when studying the convergence of numerical methods. In the literature, preservation of invariant regions has been addressed in the following special cases. In the scalar case, the existence of invariant regions corresponds to the maximum principle. On planar domains, works in this direction cover the homogeneous heat equation [10], RD scalar equations [11–13], anisotropic RD [14] and reaction–convection–diffusion scalar equations [15]. For RDSs of many equations on planar domains in the absence of cross-diffusion, the problem is addressed in [16]. On stationary surfaces, the case of RDSs of many equations in the absence of cross-diffusion is studied in [17]. The aforementioned papers consider different spatial approximation approaches. Most of them require the discretisation to be sufficiently refined, in order to preserve invariant rectangles and maximum principles. A notable exception is the lumped finite element method (LFEM) [10–12,14,15]. In this paper, we extend the results for RD systems on stationary surfaces obtained in [17] to the case when cross-diffusion is present. We propose a *fully-discrete scheme* for (1) by employing the Implicit–Explicit (IMEX) Euler scheme for the time discretisation whereby we treat implicitly the diffusion and cross-diffusion terms and explicitly the reaction terms.

The main contribution of this paper is two-fold. First, we prove sufficient conditions for the existence of invariant polytopes at the semi- and fully-discrete levels for semilinear RCDSs in (1) (i.e. in which only the kinetics are nonlinear). Second, we prove optimal error bounds for the semi- and fully-discrete schemes. We present a numerical test for the Rosenzweig–MacArthur kinetics with cross-diffusion on the unit sphere to provide an example of RCDS possessing an invariant parallelogram Σ, in which the surface FEM (SFEM) [18] in the absence of mass lumping blows-up, while the LSFEM solutions stay in the invariant region for all times. Moreover, we solve the same model with different parameters to compare the lumped and non-lumped methods for the approximation of Turing patterns on the sphere.

The present paper is structured as follows. In Section 2 we consider, for (1), the LSFEM space discretisation, the Euler IMEX/LSFEM time discretisation and, in Theorems 1 and 2, we prove sufficient conditions for the existence of invariant regions for the semi- and fully-discrete schemes, respectively. In Section 3, optimal error estimates for both the semi- and fully-discrete methods are proven in Theorems 4 and 5, respectively. Numerical tests are shown in Section 4. We conclude our work by outlining future research extensions in Section 5.

2. Reaction–cross-diffusion systems on surfaces

Let Γ be a compact, orientable, smooth surface of co-dimension one in $\mathbb{R}^3$ without boundary. We assume that Γ can be represented as the zero level set of a smooth *signed distance function* d, that is $\Gamma = \{\boldsymbol{x} \in W \mid d(\boldsymbol{x}) = 0\}$, where d is defined in an open neighbourhood W of Γ such that $\nabla d(\boldsymbol{x}) \neq \boldsymbol{0}\ \forall \boldsymbol{x} \in W$. The normal unit vector on Γ is then defined by $\boldsymbol{\nu}(\boldsymbol{x}) = \frac{\nabla d(\boldsymbol{x})}{|\nabla d(\boldsymbol{x})|}$, $\boldsymbol{x} \in \Gamma$. We assume that every point $\boldsymbol{x} \in W$ may be uniquely represented as

$$\boldsymbol{x} = \boldsymbol{a}(\boldsymbol{x}) + d(\boldsymbol{x})\boldsymbol{\nu}(\boldsymbol{a}(\boldsymbol{x})), \tag{2}$$

with $\boldsymbol{a}(\boldsymbol{x}) \in \Gamma$. A sufficient condition on the thickness of W, depending on the curvature of Γ, such that this property holds is given in [9].

For completeness' sake, we briefly recall the definitions of Sobolev and Bochner spaces on surfaces [17]. For $q \in \mathbb{N} \cup \{0\}$, the Sobolev space $H^q(\Gamma)$ is the space of functions $u : \Gamma \to \mathbb{R}$ such that, for all $i = 0,\ldots,q$, the ith order tangential derivatives, meant in a distributional sense, are $L^2(\Gamma)$, whilst $H^{-q}(\Gamma)$ is the dual space of $H^q(\Gamma)$, that is the space of linear continuous functionals on $H^q(\Gamma)$. For $p \in [1,+\infty]$, if X is a Banach space, the Bochner space $L^p([0,T];X)$ is the space of functions $u : [0,T] \to X$ such that the function $\|u\|_X : [0,T] \to \mathbb{R}$ is $L^p([0,T])$. For further details on Sobolev and Bochner spaces on surfaces we refer the interested reader to [19–21]. Given a function space S, we consider the tensor product norm on S^r defined by $\|\boldsymbol{v}\|_{S^r} := \sqrt{\sum_{i=1}^r \|v_i\|_S^2}$, for all $\boldsymbol{v} \in S^r$. For $p \in [1,+\infty]$, the $L^p([0,T];S^r)$ norms of space and time dependent functions $\boldsymbol{u} : \Gamma \times [0,T] \to \mathbb{R}^r$ are defined accordingly. Without loss of generality, we write $\|\cdot\|_S$ and $L^p([0,T];S)$ instead of $\|\cdot\|_{S^r}$ and $L^p([0,T];S^r)$, respectively. We recall that, given $n,m \in \mathbb{N}$ and any two matrices $\boldsymbol{A},\boldsymbol{B} \in \mathbb{R}^{n,m}$, the Frobenius inner product of $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined by $\boldsymbol{A} : \boldsymbol{B} := \sum_{i=1}^n \sum_{j=1}^m a_{ij}b_{ij}$. We expect that the arguments we make in the sequel still hold for systems on surfaces with boundaries in the case of homogeneous Neumann boundary conditions, i.e. zero flux on $\partial\Gamma$. However, we will confine the present analysis to the case of compact surfaces without boundary to simplify the presentation. The weak formulation of (1) is given by: find $u_1, \dots, u_r \in L^2([0,T]; H^1(\Gamma)) \cap L^\infty([0,T]\times\Gamma)$ with $\dot{u}_1, \dots, \dot{u}_r \in L^2([0,T]; H^{-1}(\Gamma))$ such that

$$\int_\Gamma \dot{u}_m \varphi_m + \sum_{k=1}^r d_{mk} \int_\Gamma \nabla_\Gamma u_k \cdot \nabla_\Gamma \varphi_m = \int_\Gamma f_m(\boldsymbol{u})\varphi_m, \qquad \forall\, \varphi_m \in H^1(\Gamma), \qquad \forall\, m = 1, \dots, r. \tag{3}$$

2.1. Space discretisation

In this section we present the necessary notations and concepts needed to formulate the finite element discretisation on stationary surfaces, following [9]. Given $h > 0$, a triangulated surface $\Gamma_h \subset W$ is defined by $\Gamma_h = \bigcup_{K \in \mathcal{K}_h} K$, where $\mathcal{K}_h$ is a set of finitely many non degenerate, non overlapping triangles, whose diameters do not exceed h and whose vertices $\{\boldsymbol{x}_i\}_{i=1}^N$ lie on Γ, such that, for $\boldsymbol{a}(\boldsymbol{x})$ as defined in (2), $\boldsymbol{a}_{|\Gamma_h}(\boldsymbol{x})$ is a one-to-one map between Γ and $\Gamma_h \subset W$.

To proceed, we define lifts and unlifts following the work in [9]. Given a function $V : \Gamma_h \to \mathbb{R}$, its *lift* $V^\ell : \Gamma \to \mathbb{R}$ is defined by $V^\ell(\boldsymbol{a}(\boldsymbol{x})) = V(\boldsymbol{x})$, $\boldsymbol{x} \in \Gamma_h$. Given a function $v : \Gamma \to \mathbb{R}$, its *unlift* $v^{-\ell} : \Gamma_h \to \mathbb{R}$ is defined by $v^{-\ell}(\boldsymbol{x}) = v(\boldsymbol{a}(\boldsymbol{x}))$, $\boldsymbol{x} \in \Gamma_h$. Next, let S_h be the space of piecewise linear functions on Γ_h defined by $S_h = \{V \in \mathcal{C}^0(\Gamma_h) \mid V_{|K} \text{ is affine } \forall K \in \mathcal{K}_h\}$ and S_h^ℓ be its lifted counterpart $S_h^\ell = \{V^\ell \mid V \in S_h\}$. Let $\{\chi_i\}_{i=1}^N$ be the nodal basis of S_h defined by $\chi_i(\boldsymbol{x}_j) = \delta_{ij}$ for all $i, j = 1, \dots, N$. For $v \in \mathcal{C}^0(\Gamma_h)$, the piecewise linear interpolant $I_h(v)$ of v is the function in S_h given by $I_h(v) = \sum_{i=1}^N v(\boldsymbol{x}_i)\chi_i$. We define the following space discretisation for the RCDS (3): find $U_1, \dots, U_r \in L^2([0,T]; S_h)$ with $\dot{U}_1, \dots, \dot{U}_r \in L^2([0,T]; S_h)$ such that

$$\int_{\Gamma_h} I_h(\dot{U}_m \varphi_m) + \sum_{k=1}^r d_{mk} \int_{\Gamma_h} \nabla_{\Gamma_h} U_k \cdot \nabla_{\Gamma_h} \varphi_m = \int_{\Gamma_h} I_h(f_m(\boldsymbol{U})\varphi_m), \quad \forall \varphi_m \in S_h, \quad \forall\, m = 1, \dots, r, \tag{4}$$

where the initial condition $(U_{0,1}, \dots, U_{0,r}) \in S_h^r$ is a suitable approximation of the initial condition $(u_{0,1}, \dots, u_{0,r})$ of the weak continuous system (3). By expressing each component U_k as

$$U_k(\boldsymbol{x}, t) = \sum_{i=1}^N \xi_{k,i}(t)\chi_i(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Gamma_h,\ t \in [0,T], \tag{5}$$

and choosing the test functions in (4) to be the nodal basis functions, we rewrite (4) as follows

$$\int_{\Gamma_h} I_h(\dot{U}_m \chi_j) + \sum_{k=1}^r d_{mk} \sum_{i=1}^N \xi_{k,i} \int_{\Gamma_h} \nabla_{\Gamma_h}\chi_i \cdot \nabla_{\Gamma_h}\chi_j = \int_{\Gamma_h} I_h(f_m(\boldsymbol{U})\chi_j), \tag{6}$$

for all $m = 1, \dots, r$ and $j = 1, \dots, N$. We define the lumped mass matrix $\overline{\boldsymbol{M}} = (\bar{m}_{ij})$ and the stiffness matrix $\boldsymbol{A} = (a_{ij})$, respectively, by

$$\bar{m}_{ij} := \int_{\Gamma_h} I_h(\chi_i\chi_j), \qquad a_{ij} := \int_{\Gamma_h} \nabla_{\Gamma_h}\chi_i \cdot \nabla_{\Gamma_h}\chi_j, \qquad i, j = 1, \dots, N.$$

We recall that the mass matrix used in the standard SFEM [9,18] is defined by $m_{ij} := \int_{\Gamma_h} \chi_i\chi_j$, for all $i, j = 1, \dots, N$. The matrix form of the LSFEM (6) is given by

$$\overline{\boldsymbol{M}}\dot{\boldsymbol{\xi}}_m + \sum_{k=1}^r d_{mk}\boldsymbol{A}\boldsymbol{\xi}_k = \overline{\boldsymbol{M}}f_m(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_r), \qquad \forall\, m = 1, \dots, r. \tag{7}$$

This system of ordinary differential equations (ODEs) (7) can also be rewritten as

$$\boldsymbol{I}_r \otimes \overline{\boldsymbol{M}} \begin{pmatrix} \dot{\boldsymbol{\xi}}_1 \\ \vdots \\ \dot{\boldsymbol{\xi}}_r \end{pmatrix} + \boldsymbol{D} \otimes \boldsymbol{A} \begin{pmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_r \end{pmatrix} = \boldsymbol{I}_r \otimes \overline{\boldsymbol{M}} \begin{pmatrix} f_1(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_r) \\ \vdots \\ f_r(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_r) \end{pmatrix}, \tag{8}$$

where $\boldsymbol{I}_r$ is the $r \times r$ identity matrix and $\otimes$ denotes the Kronecker product [22]. Since, from the properties of the Kronecker product [22], $(\boldsymbol{I}_r \otimes \overline{\boldsymbol{M}})^{-1}(\boldsymbol{D} \otimes \boldsymbol{A}) = (\boldsymbol{I}_r \otimes \overline{\boldsymbol{M}}^{-1})(\boldsymbol{D} \otimes \boldsymbol{A}) = (\boldsymbol{I}_r\boldsymbol{D}) \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A}) = \boldsymbol{D} \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A})$, we end up with the following formulation

$$\begin{pmatrix} \dot{\boldsymbol{\xi}}_1 \\ \vdots \\ \dot{\boldsymbol{\xi}}_r \end{pmatrix} = -\boldsymbol{D} \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A}) \begin{pmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_r \end{pmatrix} + \begin{pmatrix} f_1(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_r) \\ \vdots \\ f_r(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_r) \end{pmatrix}, \tag{9}$$

with $\xi_{k,i}(0) = U_{0,k}(\boldsymbol{x}_i)$ for $k = 1, \dots, r$ and $i = 1, \dots, N$. We now assume some regularity on the mesh.

Definition 1 (*Delaunay mesh*)**.** For every edge e of the triangulation $\mathcal{K}_h$, let K_1 and K_2 be the triangles sharing the edge e, α_1 and α_2 be the angles in K_1 and K_2 opposite to e, respectively. The triangulation $\mathcal{K}_h$ is said to meet the *Delaunay condition* if, for every edge e in $\mathcal{K}_h$, $\alpha_1 + \alpha_2 \leq \pi$.

The following property of Delaunay meshes was proven in [17].

Lemma 1 (*Characterisation of Delaunay meshes*)**.** *$\mathcal{K}_h$ meets the Delaunay condition if and only if $(\nabla_{\Gamma_h}\chi_i, \nabla_{\Gamma_h}\chi_j) \leq 0$ for all $i \neq j$.*

We will show that Lemma 1 plays a crucial role in the existence of invariant regions of RCDSs at the discrete levels.

2.2. Time discretisation

Applying the Euler IMEX scheme to (4) with time stepsize $\tau > 0$ and total number of timesteps given by $N_T := \left\lfloor \frac{T}{\tau} \right\rfloor$ we obtain the following fully-discrete method for (3): for all $n = 0, \dots, N_T$, for all $\varphi_1, \dots, \varphi_r \in S_h$

$$\begin{cases} \displaystyle\int_\Gamma I_h\left(\frac{U_1^{n+1} - U_1^n}{\tau}\varphi_1\right) + \sum_{k=1}^r d_{1k}\int_\Gamma \nabla_\Gamma U_k^{n+1}\cdot\nabla_\Gamma\varphi_1 = \int_\Gamma I_h(f_1(\boldsymbol{U}^n)\varphi_1), \\ \vdots \\ \displaystyle\int_\Gamma I_h\left(\frac{U_r^{n+1} - U_r^n}{\tau}\varphi_r\right) + \sum_{k=1}^r d_{rk}\int_\Gamma \nabla_\Gamma U_k^{n+1}\cdot\nabla_\Gamma\varphi_r = \int_\Gamma I_h(f_r(\boldsymbol{U}^n)\varphi_r), \end{cases} \tag{10}$$

where the initial condition $(U_1^0, \dots, U_r^0)$ coincides with that of the semi-discrete method $(U_{0,1}, \dots, U_{0,r})$. The stability estimates for (10) will rely on an energy argument. In terms of the lumped mass- and stiffness-matrices $\overline{\boldsymbol{M}}$ and $\boldsymbol{A}$ defined in Section 2.1, the scheme (10) can be written as a system of rN algebraic linear equations of the form

$$\begin{pmatrix} \xi_1^{n+1} \\ \vdots \\ \xi_r^{n+1} \end{pmatrix} = (\boldsymbol{I}_{Nr} + \tau\boldsymbol{D}\otimes(\overline{\boldsymbol{M}}^{-1}\boldsymbol{A}))^{-1}\begin{pmatrix} \xi_1^n + \tau f_1(\xi_1^n, \dots, \xi_r^n) \\ \vdots \\ \xi_r^n + \tau f_r(\xi_1^n, \dots, \xi_r^n) \end{pmatrix}, \tag{11}$$

to be solved at each timestep $t_n := n\tau$ for $n = 0, \dots, N_T$. Note that scheme (11) can be obtained equivalently by applying the IMEX Euler timestepping to the semi-discrete scheme (9). If the solutions of (9) and (11) are a-priori confined within any (possibly unbounded) set Σ contained in the domain of definition I of the kinetics and the kinetics are Lipschitz on Σ, then these solutions are well-defined at all positive times. This further motivates the study of invariant regions, addressed in the following section.

2.3. Invariant convex polytopes for the semi- and fully-discrete schemes

This section focuses on investigating an interesting property of the LSFEM discretisation of RCDSs which does not hold in the absence of lumping, that is the existence of invariant convex polytopes. For our purposes, we recall the following definition given in [21,23].

Definition 2. For the system (1), a region Σ in the phase-space $\mathbb{R}^r$ is said to be positively invariant if, whenever the initial condition $\boldsymbol{u}_0$ is in Σ, $\boldsymbol{u}$ stays in Σ as long as it exists and is unique.

Let us now consider polytopal invariant regions. Let $s \in \mathbb{N}$, let $\boldsymbol{n}^l \in \mathbb{R}^r$, $l = 1, \dots, s$ be unit vectors and let $c^l \in \mathbb{R}$, $l = 1, \dots, s$ be real constants. Let Σ be the convex polytope in the phase-space defined as the intersection of s half-hyperspaces:

$$\Sigma = \{\boldsymbol{y} \in \mathbb{R}^r \mid \boldsymbol{n}^l \cdot \boldsymbol{y} \leq c^l,\ \forall\, l = 1, \dots, s\}, \tag{12}$$

and consider its hyperfaces $\Sigma^l := \{\boldsymbol{y} \in \Sigma \mid \boldsymbol{n}^l \cdot \boldsymbol{y} = c^l\}$, $l = 1, \dots, s$. Consider the following inward flux condition for the kinetics:

$$\boldsymbol{f}(\boldsymbol{y}) \cdot \boldsymbol{n}^l(\boldsymbol{y}) < 0 \qquad \forall \boldsymbol{y} \in \Sigma^l,\ \forall\, l = 1, \dots, s, \tag{13}$$

and the following compatibility condition between Σ and $\boldsymbol{D}$

$$\boldsymbol{n}^l \text{ is a left eigenvector of } \boldsymbol{D} \qquad \forall\, l = 1, \dots, s. \tag{14}$$

In order for the region Σ to be invariant, (i) condition (13) is sufficient in the absence of cross-diffusion when Γ is a Riemannian manifold without boundary [21], while (ii) conditions (13) and (14) are sufficient in the presence of cross-diffusion when Γ is a k-dimensional domain in $\mathbb{R}^k$, $k \in \mathbb{N}$ [23]. In the following theorems we prove that, in the presence of cross-diffusion on a compact surface, under assumptions (13) and (14), Σ is an invariant region for the semi- (9) and fully-discrete (11) systems conditionally on τ.

Theorem 1 (*Invariant convex polytopes for the semi-discrete system (9)*). *Let the kinetics $\boldsymbol{f}$ be Lipschitz on the polytope Σ in (12) and assume that (13)-(14) hold. Then Σ is an invariant region for the semi-discrete problem (9).*

Proof. It suffices to prove that the rN-dimensional polytope $\bar{\Sigma} = \Sigma^N$ is an invariant region for the ODE system (9), i.e. we have to prove that the vector field on the right-hand-side of (9), computed on the boundary of $\bar{\Sigma}$, points towards the interior of $\bar{\Sigma}$. To this end, let $(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r)^T$ be a point on $\partial\bar{\Sigma}$. This means that there exist $i = 1, \ldots, N$ and $l = 1, \ldots, s$ such that $\boldsymbol{n}^l \cdot \xi_{:,i} = c^l$. Then, $\boldsymbol{n}^l \cdot \xi_{:,j} \leq c^l$ for $j \neq i$, that implies

$$\xi_{:,j} \cdot \boldsymbol{n}^l \leq \xi_{:,i} \cdot \boldsymbol{n}^l, \quad j \neq i. \tag{15}$$

All we have to prove is that $\dot{\xi}_{:,i} \cdot \boldsymbol{n}^l$ is nonpositive. But since, from (13), $\boldsymbol{f}(\xi_{1,i}, \ldots, \xi_{r,i}) \cdot \boldsymbol{n}^l < 0$, it remains to prove that

$$-\left(\boldsymbol{D} \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A})\begin{pmatrix}\boldsymbol{\xi}_1\\ \vdots \\ \boldsymbol{\xi}_r\end{pmatrix}\right)_{:,i} \cdot \boldsymbol{n}^l \leq 0. \tag{16}$$

Since $\boldsymbol{n}^l$ is a left eigenvector of $\boldsymbol{D}$ (with eigenvalue $\lambda^l > 0$), the left-hand side of (16) is equal to

$$\begin{aligned}
-\lambda^l\left(\boldsymbol{I}_r \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A})\begin{pmatrix}\boldsymbol{\xi}_1\\ \vdots \\ \boldsymbol{\xi}_r\end{pmatrix}\right)_{:,i} \cdot \boldsymbol{n}^l &= -\lambda^l \sum_{k=1}^{r} n_k^l\left(\overline{\boldsymbol{M}}^{-1}\boldsymbol{A}\boldsymbol{\xi}_k\right)_i = -\lambda^l \sum_{k=1}^{r} n_k^l \bar{m}_{ii}^{-1} \sum_{j=1}^{N} a_{ij}\xi_{k,j} \\
&= -\lambda^l \bar{m}_{ii}^{-1} \sum_{j=1}^{N} a_{ij} \sum_{k=1}^{r} n_k^l \xi_{k,j} = -\lambda^l \bar{m}_{ii}^{-1} \sum_{j=1}^{N} a_{ij}(\xi_{:,j} \cdot \boldsymbol{n}^l) \\
&= -\lambda^l \bar{m}_{ii}^{-1}\left(a_{ii}(\xi_{:,i} \cdot \boldsymbol{n}^l) + \sum_{j \in \{1,\ldots,N\}\setminus\{i\}} a_{ij}(\xi_{:,j} \cdot \boldsymbol{n}^l)\right).
\end{aligned} \tag{17}$$

From $a_{ij} \leq 0$, $i \neq j$ (Lemma 1) and (15), the right-hand side of (17) is less than or equal to

$$\lambda^l \bar{m}_{ii}^{-1}(\xi_{:,i} \cdot \boldsymbol{n}^l)\left(-a_{ii} + \sum_{j \in \{1,\ldots,N\}\setminus\{i\}} (-a_{ij})\right) = -\lambda^l \bar{m}_{ii}^{-1}(\xi_{:,i} \cdot \boldsymbol{n}^l) \sum_{j=1}^{N} a_{ij}. \tag{18}$$

From the definition of $\boldsymbol{A}$ we have

$$\sum_{j=1}^{N} a_{ij} = \int_{\Gamma_h} \nabla_{\Gamma_h}\chi_i \cdot \nabla_{\Gamma_h} \sum_{j=1}^{N} \chi_i. \tag{19}$$

Since Γ_h has no boundary, $\sum_{j=1}^{N} \chi_i(\boldsymbol{x}) = 1 \; \forall\, \boldsymbol{x} \in \Gamma_h$ and thus

$$\nabla_{\Gamma_h} \sum_{j=1}^{N} \chi_i(\boldsymbol{x}) = 0, \qquad \forall\, \boldsymbol{x} \in \Gamma_h. \tag{20}$$

By combining (17)– (20), we have proven (16), which completes the proof. □

For the proof of the following theorem, we need two lemmas whose statements and proofs are given in the Appendix. In particular, Lemma A extends Lemma 15.5 in [24].

Theorem 2 (*Invariant convex polytopes for the fully-discrete scheme (11)*). *Let the kinetics $\boldsymbol{f}$ be Lipschitz on the polytope Σ in (12) and assume that (13)-(14) hold. Then Σ is an invariant region for the fully-discrete problem (11) if the timestep τ fulfils*

$$\tau \leq \bar{\tau} := \frac{1}{\max_{l=1,\ldots,s}\sqrt{\sum_{k=1}^{r}(n_k^l L_k)^2}}, \tag{21}$$

where $L_1, \ldots, L_r$ are the Lipschitz constant of the kinetics $f_1, \ldots, f_r$, respectively.

Proof. Given $n = 0, \ldots, N_T$ and $\boldsymbol{U}^n \in \Sigma$, we must ensure that $\boldsymbol{U}^{n+1} \in \Sigma$, i.e., that it satisfies $\boldsymbol{n}^l \cdot \boldsymbol{U}^{n+1} \leq c^l$, $l = 1, \ldots, s$. Since $\boldsymbol{U}^{n+1}$ is an S_h function, it suffices to verify that $\boldsymbol{U}^{n+1}$ satisfies the inequality at the gridpoints. Using the definition of the fully-discrete scheme (11) we wish to show that

$$\boldsymbol{n}^l \cdot \boldsymbol{U}^{n+1} = (\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)\mathcal{M}(\boldsymbol{\xi}^n + \tau\boldsymbol{f}^n) \leq c^l \boldsymbol{1}_N, \qquad l = 1, \ldots, s, \tag{22}$$

where $\mathcal{M} := (\boldsymbol{I}_{rN} + \tau \boldsymbol{D} \otimes \overline{\boldsymbol{M}}^{-1} \boldsymbol{A})^{-1}$ and $\boldsymbol{f}^n := \boldsymbol{f}(\boldsymbol{\xi}_1^n, \dots, \boldsymbol{\xi}_r^n)$. If $\boldsymbol{v} \in \mathbb{R}^r$ is such that $\boldsymbol{n}^l \cdot \boldsymbol{v} = c^l$, (22) becomes

$$(\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)\mathcal{M}(\boldsymbol{\xi}^n + \tau \boldsymbol{f}^n) \le (\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)(\boldsymbol{v} \otimes \boldsymbol{I}_N), \quad l = 1, \dots, s. \tag{23}$$

By applying Lemma B in Appendix to the right-hand side of (23), we end up with

$$(\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)\mathcal{M}(\boldsymbol{\xi}^n + \tau \boldsymbol{f}^n - \boldsymbol{v} \otimes \boldsymbol{1}_N) \le \boldsymbol{0}, \quad l = 1, \dots, s. \tag{24}$$

From Lemma A in Appendix, it suffices to prove that

$$(\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)(\boldsymbol{\xi}^n + \tau \boldsymbol{f}^n - \boldsymbol{v} \otimes \boldsymbol{1}_N) \le \boldsymbol{0}, \quad l = 1, \dots, s, \tag{25}$$

but, recalling that $\boldsymbol{n}^l \cdot \boldsymbol{v} = c^l$, (25) is equivalent to

$$(\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)(\boldsymbol{\xi}^n + \tau \boldsymbol{f}^n) \le c^l \boldsymbol{1}_N, \quad l = 1, \dots, s. \tag{26}$$

We now observe that $\boldsymbol{d}^{l,n} := c^l \boldsymbol{1}_N - (\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)\boldsymbol{\xi}^n$ is a vector in $\mathbb{R}^N$ such that, for all $i = 1, \dots, N$, the ith component $d_i^{l,n} = c^l - \boldsymbol{n}^l \cdot \boldsymbol{\xi}^n_{:,i}$ is the oriented distance between the solution $\boldsymbol{\xi}^n_{:,i}$ on the ith nodal point and the hyperplane Σ^l. We then rewrite (26) as $\tau(\boldsymbol{n}^{l,T} \otimes \boldsymbol{I}_N)\boldsymbol{f}^n \le \boldsymbol{d}^{l,n}$, $l = 1, \dots, s$. Componentwise, we have $\tau \boldsymbol{n}^l \cdot \boldsymbol{f}^n_{:,i} \le d_i^{l,n}$, $i = 1, \dots, N$, $l = 1, \dots, s$. Now, since $\boldsymbol{U}^n \in \Sigma$, we can upper-bound this last inequality in terms of the oriented distances $\boldsymbol{d}^{l,n}$ and the directional Lipschitz constant $\tilde{L}^l$ of the kinetics $\boldsymbol{f}$ along the outward normal $\boldsymbol{n}^l$, obtaining $\tau \tilde{L}^l d_i^{l,n} \le d_i^{l,n}$, $i = 1, \dots, N$, $l = 1, \dots, s$, but since $\tilde{L}^l \le \sqrt{\sum_{k=1}^r (n_k^l)^2 L_k^2}$, the result follows. □

3. Stability and error analysis

Next we prove in this section stability estimates and optimal $L^\infty([0,T], L^2(\Gamma))$ error bounds for the semi-discrete (9) and the fully-discrete (11) solutions of the RCDS (1) of $r \in \mathbb{N}$ equations. First, we proceed to recall some preliminaries and basic notations.

The lumped L^2 product (see for instance [11,24–26]) defined by $(U,V)_h := \int_{\Gamma_h} I_h(UV)$, $U, V \in L^2(\Gamma_h)$, where I_h is the piecewise linear interpolant defined in Section 2.1, induces the norm $\|U\|_h := \sqrt{(U,U)_h}$, $U \in S_h$, which is equivalent to $\|\cdot\|_{L^2(\Gamma_h)}$, uniformly with respect to h (see [27] for the proof):

$$\|U\|_{L^2(\Gamma_h)} \le \|U\|_h \le C\|U\|_{L^2(\Gamma_h)}, \quad U \in S_h, \quad h > 0. \tag{27}$$

Let us define the "broken" Sobolev space $H_h^2(\Gamma_h) := H^1(\Gamma_h) \cap \prod_{K \in \mathcal{K}_h} H^2(K)$, endowed with the norm $\|U\|^2_{H_h^2(\Gamma_h)} := \sum_{K \in \mathcal{K}_h} \|U\|^2_{H^2(K)}$, $U \in H_h^2(\Gamma_h)$. For the error $\varepsilon_h(U,V) := \int_{\Gamma_h}(UV - I_h(UV))$ in the lumped quadrature rule $(U,V)_h$, if $U \in H_h^2(\Gamma_h)$ and $V \in S_h$, then the following estimate holds (see [11]):

$$|\varepsilon_h(U,V)| \le ch^2 \|U\|_{H_h^2(\Gamma_h)} \|V\|_{H^1(\Gamma_h)}. \tag{28}$$

We remark that for the case of RDs without cross-diffusion, inequalities (27) and (28) have been proven on planar triangulations in [26] and [11], respectively. By using an affine map argument, these inequalities can be easily extended to triangulated surfaces since their respective proofs are done piecewise on each triangle. The following equivalences between the norms of a function U defined on Γ_h and its lifted counterpart U^ℓ can be found in [9].

Lemma 2 (*Equivalence of element-wise norms under lifting,* [9]). *Let $K \in \mathcal{K}_h$, $\tilde{K} := \boldsymbol{a}(T) \subset \Gamma$, where the map $\boldsymbol{a}(\boldsymbol{x})$ is given in (2), and $U : K \to \mathbb{R}$. If the norms exist, then the following inequalities hold*

$$c\|U\|_{L^2(K)} \le \|U^\ell\|_{L^2(\tilde{K})} \le C\|U\|_{L^2(K)}; \tag{29}$$

$$c\|\nabla_T U\|_{L^2(K)} \le \|\nabla_{\tilde{K}} U^\ell\|_{L^2(\tilde{K})} \le C\|\nabla_K U\|_{L^2(K)}. \tag{30}$$

From the previous lemma we derive the following estimate for the broken H^2 norm of U.

Lemma 3 (*Dominance of $H^2(\Gamma)$ norm over $H^2(\Gamma_h)$ norm*). *If $u \in H^2(\Gamma)$, then $u^{-\ell} \in H_h^2(\Gamma_h)$ and*

$$\|u^{-\ell}\|_{H_h^2(\Gamma_h)} \le C\|u\|_{H^2(\Gamma)}. \tag{31}$$

Proof. The reader is referred to consult [17] for the proof. □

When lifting integrals, a geometric error must be taken into account. The following equality holds (see [9, p. 317])

$$\int_{\Gamma_h} UV = \int_\Gamma \frac{U^\ell V^\ell}{\delta_h^\ell}, \quad \forall\, U, V \in L^2(\Gamma_h), \tag{32}$$

where the function $\delta_h^\ell : \Gamma \to \mathbb{R}$, defined in [9, p. 310], fulfils

$$\left\| 1 - \frac{1}{\delta_h^\ell} \right\|_{L^\infty(\Gamma)} \leq Ch^2. \tag{33}$$

For the following proofs we need to define, for any positive definite matrix $\boldsymbol{B} \in \mathbb{R}^r$, the seminorm $|\cdot|_{\boldsymbol{B},h}$ on $(H^1(\Gamma_h))^r$ by

$$|\boldsymbol{U}|^2_{\boldsymbol{B},h} := \int_{\Gamma_h} \boldsymbol{B}\nabla_{\Gamma_h}\boldsymbol{U} : \nabla_{\Gamma_h}\boldsymbol{U} = \int_{\Gamma_h} \boldsymbol{B}^s\nabla_{\Gamma_h}\boldsymbol{U} : \nabla_{\Gamma_h}\boldsymbol{U}, \qquad \forall\, \boldsymbol{U} \in H^1(\Gamma_h)^r, \tag{34}$$

where $\boldsymbol{B}^s := \frac{\boldsymbol{B}+\boldsymbol{B}^T}{2}$ is the symmetric part of $\boldsymbol{B}$. It is well-known that a matrix is positive definite if and only if its symmetric part is positive definite. Then, the eigenvalues λ_i, $i = 1, \ldots, r$, of $\boldsymbol{B}^s$ are real and positive. It follows that

$$\min_{i=1,\ldots,r} (\lambda_i)|\boldsymbol{U}|^2_{H^1(\Gamma_h)} \leq |\boldsymbol{U}|^2_{\boldsymbol{B},h} \leq \max_{i=1,\ldots,r} (\lambda_i)|\boldsymbol{U}|^2_{H^1(\Gamma_h)}, \quad \forall \boldsymbol{U} \in (H^1(\Gamma_h))^r, \tag{35}$$

i.e. the seminorm (34) is equivalent to $|\cdot|_{H^1(\Gamma_h)}$.

We employ the usual energy argument techniques to carry out the following stability estimates. Note that due to the existence of an invariant region, these estimates will not depend exponentially on time since they will not rely on Grönwall's lemma. The only requirement is that the reaction kinetics $\boldsymbol{f}$ in (1) are Lipschitz locally in the invariant region and not globally Lipschitz.

Lemma 4 (*Stability estimates for the weak formulation (3)*). *If $\boldsymbol{u}$ is the solution of (3), Σ as in (12) is a bounded invariant region for (3), $\boldsymbol{f}$ is Lipschitz (and thus bounded) on Σ and $\boldsymbol{u}_0 \in \Sigma$, then the following estimates hold*

$$\sup_{t\in[0,T]} \|\boldsymbol{u}\|^2_{L^2(\Gamma)} + \int_0^T \|\nabla_\Gamma \boldsymbol{u}\|^2_{L^2(\Gamma)} \leq C\left(T + \|\boldsymbol{u}_0\|^2_{L^2(\Gamma)}\right), \tag{36}$$

$$\int_0^T \|\dot{\boldsymbol{u}}\|^2_{L^2(\Gamma)} + \sup_{t\in[0,T]} \|\nabla_\Gamma \boldsymbol{u}\|^2_{L^2(\Gamma)} \leq C\left(T + \|\nabla_\Gamma \boldsymbol{u}_0\|^2_{L^2(\Gamma)}\right), \tag{37}$$

for all $T > 0$, where C is a constant independent of T and $\boldsymbol{u}_0$.

Proof. The proof relies on the usual energy arguments, see for instance [18]. □

In the next lemmas we show analogous estimates for the semi- and fully-discrete problems.

Lemma 5 (*Stability estimates for the semi-discrete system (4)*). *If $\boldsymbol{U}$ is the solution of (4), Σ is a bounded invariant region for (4), $\boldsymbol{f}$ is Lipschitz on Σ and $\boldsymbol{U}_0 \in \Sigma$, then*

$$\sup_{t\in[0,T]} \|\boldsymbol{U}\|^2_{L^2(\Gamma_h)} + \int_0^T \|\nabla_\Gamma \boldsymbol{U}\|^2_{L^2(\Gamma_h)} \leq C\left(T + \|\boldsymbol{U}_0\|^2_{L^2(\Gamma_h)}\right), \tag{38}$$

$$\int_0^T \|\dot{\boldsymbol{U}}\|^2_{L^2(\Gamma_h)} + \sup_{t\in[0,T]} \|\nabla_\Gamma \boldsymbol{U}\|^2_{L^2(\Gamma_h)} \leq C\left(T + \|\nabla_\Gamma \boldsymbol{U}_0\|^2_{L^2(\Gamma_h)}\right), \tag{39}$$

for all $T > 0$, where C is a constant independent of T and $\boldsymbol{U}_0$.

Proof. We use an energy argument as in the previous lemma and then use the equivalence (27) between the norms $\|\cdot\|_h$ and $\|\cdot\|_{L^2(\Gamma_h)}$. □

Lemma 6 (*Stability estimates for the fully-discrete system (10)*). *Let $\tau > 0$. If $\boldsymbol{U}^n$, $n = 0, \ldots, N_T$, is the solution of (10), Σ is a bounded invariant region for (10), $\boldsymbol{f}$ is Lipschitz on Σ and $\boldsymbol{U}^0 \in \Sigma$, then*

$$\|\boldsymbol{U}^{m+1}\|^2_{L^2(\Gamma_h)} + \tau \sum_{n=0}^{m} \|\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|^2_{L^2(\Gamma_h)} \leq C(\|\boldsymbol{U}^0\|_{L^2(\Gamma_h)} + T), \tag{40}$$

$$\frac{1}{\tau}\sum_{n=0}^{m} \|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|^2_{L^2(\Gamma_h)} + \|\nabla_{\Gamma_h}\boldsymbol{U}^{m+1}\|^2_{L^2(\Gamma_h)} \leq C(\|\nabla_{\Gamma_h}\boldsymbol{U}^0\|^2_{L^2(\Gamma_h)} + T), \tag{41}$$

for all $m = 0, \ldots, N_T$ and $T > 0$, where C is a constant independent of T and $\boldsymbol{U}^0$.

Proof. By summing over the equations in (10) and choosing $\boldsymbol{\phi}^n = \boldsymbol{U}^{n+1}$ we have

$$\frac{1}{\tau}\left(\|\boldsymbol{U}^{n+1}\|_h^2 - \int_{\Gamma_h} I_h(\boldsymbol{U}^n : \boldsymbol{U}^{n+1})\right) + |\boldsymbol{U}^{n+1}|^2_{\boldsymbol{D},h} = \int_{\Gamma_h} I_h(\boldsymbol{f}(\boldsymbol{U}^n) : \boldsymbol{U}^{n+1}).$$

After multiplying by τ, Cauchy–Schwarz inequality yields

$$\|\boldsymbol{U}^{n+1}\|_h^2 + \tau|\boldsymbol{U}^{n+1}|_{\boldsymbol{D},h} \leq \|\boldsymbol{U}^{n+1}\|_h\|\boldsymbol{U}^n\|_h + \tau\|\boldsymbol{f}(\boldsymbol{U})^n\|_h\|\boldsymbol{U}^{n+1}\|_h.$$

Since $\boldsymbol{U}^n$ and $\boldsymbol{U}^{n+1} \in \Sigma$ and $\boldsymbol{f}$ is Lipschitz on Σ, the last term on the right-hand side is bounded by some constant $C > 0$: $\|\boldsymbol{U}^{n+1}\|_h^2 + \tau|\boldsymbol{U}^{n+1}|_{\boldsymbol{D},h} \leq \|\boldsymbol{U}^{n+1}\|_h\|\boldsymbol{U}^n\|_h + C\tau$. Young's inequality yields $\|\boldsymbol{U}^{n+1}\|_h^2 + \tau|\boldsymbol{U}^{n+1}|_{\boldsymbol{D},h}^2 \leq \|\boldsymbol{U}^n\|_h^2 + C\tau$. We sum for $n = 0, \ldots, m$ to obtain

$$\|\boldsymbol{U}^{m+1}\|_h^2 + \tau\sum_{n=0}^{m}|\boldsymbol{U}^{n+1}|_{\boldsymbol{D},h}^2 \leq \|\boldsymbol{U}^0\|_h^2 + Cm\tau.$$

By using (27) and (35) and $m \leq N_T$, (40) follows immediately. Summing over the equations in (10) and choosing $\boldsymbol{\phi}^n = \boldsymbol{D}(\boldsymbol{U}^{n+1} - \boldsymbol{U}^n)$, since $\boldsymbol{D}$ is constant and positive definite we have

$$\frac{1}{\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 - \int_{\Gamma_h}\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1} : \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^n \leq C\int_{\Gamma_h} I_h(\boldsymbol{f}(\boldsymbol{U}^n) : \boldsymbol{D}(\boldsymbol{U}^{n+1} - \boldsymbol{U}^n)).$$

Cauchy–Schwarz inequality yields

$$\begin{aligned}
&\frac{1}{\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 \\
&\qquad \leq \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}\|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^n\|_{L^2(\Gamma_h)} + C\|\boldsymbol{f}(\boldsymbol{U}^n)\|_h\|\boldsymbol{D}(\boldsymbol{U}^{n+1} - \boldsymbol{U}^n)\|_h.
\end{aligned}$$

Since $\boldsymbol{f}$ is Lipschitz -and thus bounded on Σ, say $\max_\Sigma|\boldsymbol{f}| = C$, we can bound the last term on the right-hand side as follows:

$$\frac{1}{\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 \leq \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}\|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^n\|_{L^2(\Gamma_h)} + C\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h.$$

Young's inequality yields

$$\begin{aligned}
&\frac{1}{\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 \\
&\qquad \leq \frac{1}{2}\left(\|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^n\|_{L^2(\Gamma_h)}^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2\right) + C\tau + \frac{1}{2\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2.
\end{aligned}$$

Rearranging terms and multiplying by 2 we have

$$\frac{1}{\tau}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 \leq \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{n+1}\|_{L^2(\Gamma_h)}^2 + C\tau. \tag{42}$$

By summing (42) for $n = 0, \ldots, m$ we have

$$\frac{1}{\tau}\sum_{n=0}^{m}\|\boldsymbol{U}^{n+1} - \boldsymbol{U}^n\|_h^2 + \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^{m+1}\|_{L^2(\Gamma_h)}^2 \leq \|\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U}^0\|_{L^2(\Gamma_h)}^2 + Cm\tau.$$

Now, since $\boldsymbol{D}$ is positive definite, by using (27) and $m \leq N_T$, (41) finally follows. □

In what follows, we adopt the surface Ritz projection considered in [28–30] to prove the convergence of the semi- and fully-discrete methods.

Definition 3 (*Ritz projection*)**.** Given $u : [0, T] \to H^1(\Gamma)$, the Ritz projection of u is the unique function $\bar{U} : [0, T] \to S_h$ such that

$$\int_{\Gamma_h}\nabla_{\Gamma_h}\bar{U}\cdot\nabla_{\Gamma_h}\varphi = \int_\Gamma\nabla_\Gamma u\cdot\nabla_\Gamma\varphi^\ell \quad \text{and} \quad \int_{\Gamma_h}\bar{U} = \int_\Gamma u, \qquad \varphi \in S_h. \tag{43}$$

We note that this definition is different from the one considered in [31]. The following error estimates for the Ritz projection can be found in [28,29].

Theorem 3 (*Error estimates for the Ritz projection*)**.** *Given $u : [0, T] \to H^2(\Gamma)$ such that $\dot{u} : [0, T] \to H^2(\Gamma)$, the error in the Ritz projection satisfies the following bounds*

$$\|u - \bar{U}^\ell\|_{L^2(\Gamma)} + h\|\nabla_\Gamma(u - \bar{U}^\ell)\|_{L^2(\Gamma)} \leq Ch^2\|u\|_{H^2(\Gamma)}, \tag{44}$$

$$\|\dot{u} - \dot{\bar{U}}^\ell\|_{L^2(\Gamma)} + h\|\nabla_\Gamma(\dot{u} - \dot{\bar{U}}^\ell)\|_{L^2(\Gamma)} \leq Ch^2\left(\|u\|_{H^2(\Gamma)} + \|\dot{u}\|_{H^2(\Gamma)}\right). \tag{45}$$

From here onwards, we will denote by $\bar{\boldsymbol{U}}$ the componentwise Ritz projection of a given vector function $\boldsymbol{u}$. This entails that the estimates (44)–(45) still hold in their respective tensor product norms.

Theorem 4 (*Error estimate for the semi-discrete solution (4)*). *Assume that Σ is a bounded invariant region for (3) and (4), that $\boldsymbol{f} \in \mathcal{C}^2(\Sigma)$ and that $\boldsymbol{u}_0, \boldsymbol{U}_0 \in \Sigma$. If the solution $\boldsymbol{u}$ of (3) and its time derivative $\dot{\boldsymbol{u}}$ are $L^\infty([0,T]; H^2(\Gamma))$ and $\|\boldsymbol{u}_0 - \boldsymbol{U}_0^\ell\|_{L^2(\Gamma)} \le Ch^2$, then the following estimate holds*

$$\|\boldsymbol{u} - \boldsymbol{U}^\ell\|_{L^2(\Gamma)} \le C(\boldsymbol{u}, T)h^2,$$

where $C(\boldsymbol{u}, T)$ is a constant depending on $\boldsymbol{u}$ and T.

Proof. Following [17, Theorem 7], let us write the error as $\boldsymbol{U}^\ell - \boldsymbol{u} = (\boldsymbol{U}^\ell - \bar{\boldsymbol{U}}^\ell) + (\bar{\boldsymbol{U}}^\ell - \boldsymbol{u}) =: \boldsymbol{\theta}^\ell + \boldsymbol{\rho}^\ell$. Since $\boldsymbol{u}$ and $\dot{\boldsymbol{u}}$ are $L^\infty([0,T], H^2(\Gamma))$, from the error estimates (44)– (45) for the Ritz projection and (29)– (30) we have that

$$\|\boldsymbol{\rho}\|_{L^2(\Gamma_h)} \le C\|\boldsymbol{\rho}^\ell\|_{L^2(\Gamma)} = C\|\bar{\boldsymbol{U}}^\ell - \boldsymbol{u}\|_{L^2(\Gamma)} \le Ch^2\,\|\boldsymbol{u}\|_{H^2(\Gamma)}, \tag{46}$$

$$\|\dot{\boldsymbol{\rho}}\|_{L^2(\Gamma_h)} + h\|\nabla_{\Gamma_h}\dot{\boldsymbol{\rho}}\|_{L^2(\Gamma_h)} \le Ch^2\left(\|\boldsymbol{u}\|_{H^2(\Gamma)} + \|\dot{\boldsymbol{u}}\|_{H^2(\Gamma)}\right). \tag{47}$$

It remains to show the convergence for $\boldsymbol{\theta}^\ell$. For the sake of simplicity, we derive an estimate for $\boldsymbol{\theta}$ in the norm $\|\cdot\|_h$ and then we will use (27) and (29) to estimate $\|\boldsymbol{\theta}^\ell\|_{L^2(\Gamma)}$. In the weak and semi-discrete formulations (3) and (4) we choose the same test functions φ_m, $m = 1,\dots,r$, under lifting. By subtracting these two formulations and summing over $m = 1,\dots,r$, we have

$$\begin{aligned}&\left(\int_\Gamma \dot{\boldsymbol{u}} : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} I_h(\dot{\boldsymbol{U}} : \boldsymbol{\varphi})\right) + \left(\int_\Gamma \boldsymbol{D}\nabla_\Gamma \boldsymbol{u} : \nabla_\Gamma \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U} : \nabla_{\Gamma_h}\boldsymbol{\varphi}\right)\\ &\quad = \left(\int_\Gamma \boldsymbol{f}(\boldsymbol{u}) : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} I_h(\boldsymbol{f}(\boldsymbol{U}) : \boldsymbol{\varphi})\right).\end{aligned} \tag{48}$$

Using (32) and (43) we rearrange the terms between brackets in (48) as follows

$$\int_\Gamma \dot{\boldsymbol{u}} : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} I_h(\dot{\boldsymbol{U}} : \boldsymbol{\varphi}) = \int_\Gamma \left(1 - \frac{1}{\delta_h^\ell}\right)\dot{\boldsymbol{u}} : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} \dot{\boldsymbol{\rho}} : \boldsymbol{\varphi} + \varepsilon_h(\dot{\bar{\boldsymbol{U}}}, \boldsymbol{\varphi}) - \int_{\Gamma_h} I_h(\dot{\boldsymbol{\theta}} : \boldsymbol{\varphi}),$$

$$\begin{aligned}\int_\Gamma \boldsymbol{D}\nabla_\Gamma \boldsymbol{u} : \nabla_\Gamma \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U} : \nabla_{\Gamma_h}\boldsymbol{\varphi} &= \int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\bar{\boldsymbol{U}} : \nabla_{\Gamma_h}\boldsymbol{\varphi} - \int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{U} : \nabla_{\Gamma_h}\boldsymbol{\varphi}\\ &= -\int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{\theta} : \nabla_{\Gamma_h}\boldsymbol{\varphi},\end{aligned}$$

$$\begin{aligned}\int_\Gamma \boldsymbol{f}(\boldsymbol{u}) : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} I_h(\boldsymbol{f}(\boldsymbol{U}) : \boldsymbol{\varphi}) &= \int_\Gamma \boldsymbol{f}(\boldsymbol{u}) : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} \boldsymbol{f}(\boldsymbol{u}^{-\ell}) : \boldsymbol{\varphi} + \int_{\Gamma_h} \boldsymbol{f}(\boldsymbol{u}^{-\ell}) : \boldsymbol{\varphi}\\ &\quad - \int_{\Gamma_h} I_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell}) : \boldsymbol{\varphi}) + \int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{u}^{-\ell}) - \boldsymbol{f}(\boldsymbol{U})) : \boldsymbol{\varphi})\\ &= \int_\Gamma \left(1 - \frac{1}{\delta_h^\ell}\right)\boldsymbol{f}(\boldsymbol{u}) : \boldsymbol{\varphi}^\ell + \varepsilon_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell}), \boldsymbol{\varphi}) + \int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{u}^{-\ell}) - \boldsymbol{f}(\boldsymbol{U})) : \boldsymbol{\varphi}).\end{aligned}$$

By using these relations in (48) we obtain

$$\begin{aligned}\int_{\Gamma_h} I_h(\dot{\boldsymbol{\theta}} : \boldsymbol{\varphi}) + \int_{\Gamma_h} \boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{\theta} : \nabla_{\Gamma_h}\boldsymbol{\varphi} &= \int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{U}) - \boldsymbol{f}(\boldsymbol{u}^{-\ell})) : \boldsymbol{\varphi}) - \varepsilon_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell}), \boldsymbol{\varphi})\\ &\quad - \int_\Gamma \left(1 - \frac{1}{\delta_h^\ell}\right)\boldsymbol{f}(\boldsymbol{u}) : \boldsymbol{\varphi}^\ell - \int_{\Gamma_h} \dot{\boldsymbol{\rho}} : \boldsymbol{\varphi} + \varepsilon_h(\dot{\bar{\boldsymbol{U}}}, \boldsymbol{\varphi}) + \int_\Gamma \left(1 - \frac{1}{\delta_h^\ell}\right)\dot{\boldsymbol{u}} : \boldsymbol{\varphi}^\ell.\end{aligned} \tag{49}$$

In (49) we choose $\boldsymbol{\varphi} = \boldsymbol{\theta}$. For the first term of (49) we observe that

$$\int_{\Gamma_h} I_h(\dot{\boldsymbol{\theta}} : \boldsymbol{\theta}) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\boldsymbol{\theta}\|_h^2. \tag{50}$$

We estimate the single terms on the right-hand side of (49) in turn. By using the Cauchy–Schwarz inequality, the Lipschitz continuity of $\boldsymbol{f}$, the definition of $\boldsymbol{\theta}$, (27), (29) and (46), we have that

$$\begin{aligned}\left|\int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{U}) - \boldsymbol{f}(\boldsymbol{u}^{-\ell})) : \boldsymbol{\theta})\right| \le C\|\boldsymbol{U} - \boldsymbol{u}^{-\ell}\|_h\,\|\boldsymbol{\theta}\|_h &\le C\left(\|\boldsymbol{\rho}\|_{L^2(\Gamma)} + \|\boldsymbol{\theta}\|_h\right)\|\boldsymbol{\theta}\|_h\\ &= C(\boldsymbol{u})(h^2 + \|\boldsymbol{\theta}\|_h)\,\|\boldsymbol{\theta}\|_h.\end{aligned} \tag{51}$$

By using the estimate (28) for ε_h, (31), the regularity assumptions $\boldsymbol{f} \in \mathcal{C}^2(\Sigma)$ and $\boldsymbol{u} \in L^\infty([0,T], H^2(\Gamma))$, and by applying the chain rule to the composite function $\boldsymbol{f}(\boldsymbol{u})$ it follows that

$$\begin{aligned}&\left|\varepsilon_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell}), \boldsymbol{\theta})\right| \le Ch^2\|\boldsymbol{f}(\boldsymbol{u}^{-\ell})\|_{H_h^2(\Gamma_h)}\,\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)} \le Ch^2\|\boldsymbol{f}(\boldsymbol{u})\|_{H^2(\Gamma)}\,\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)}\\ &\quad \le Ch^2\|\boldsymbol{f}\|_{\mathcal{C}^2(\Sigma)}\|\boldsymbol{u}\|_{H^2(\Gamma)}\,\|\boldsymbol{\theta}\|_{H^1(\Gamma)} \le Ch^2\,\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)}.\end{aligned} \tag{52}$$

Since $\boldsymbol{f}$ is Lipschitz over the compact region Σ, then $\boldsymbol{f} \in L^\infty(\Sigma)$. Hence, by using the Cauchy–Schwarz inequality, (29) and the geometric estimate (33) we have

$$\left|\int_\Gamma \left(1-\frac{1}{\delta_h^\ell}\right)\boldsymbol{f}(\boldsymbol{u}):\boldsymbol{\theta}^\ell\right| \le \left\|1-\frac{1}{\delta_h^\ell}\right\|_{L^\infty(\Gamma)} \|\boldsymbol{f}(\boldsymbol{u})\|_{L^2(\Gamma)} \left\|\boldsymbol{\theta}^\ell\right\|_{L^2(\Gamma)} \le Ch^2 \|\boldsymbol{\theta}\|_{L^2(\Gamma_h)}. \tag{53}$$

From the Cauchy–Schwarz inequality, the error estimate (47) for $\dot{\boldsymbol{\rho}}$,(29) we have

$$\left|\int_{\Gamma_h} \dot{\boldsymbol{\rho}}:\boldsymbol{\theta}\right| \le C\,\|\dot{\boldsymbol{\rho}}\|_{L^2(\Gamma_h)} \|\boldsymbol{\theta}\|_{L^2(\Gamma_h)} \le C(\boldsymbol{u})h^2 \|\boldsymbol{\theta}\|_{L^2(\Gamma_h)}. \tag{54}$$

From the estimate (28) for ε_h, the estimate (47) for $\dot{\boldsymbol{\rho}}$, (29) and (30), the triangle inequality and $\dot{\bar{\boldsymbol{U}}}, \boldsymbol{\theta} \in S_h$ we have

$$\begin{aligned}\left|\varepsilon_h(\dot{\bar{\boldsymbol{U}}},\boldsymbol{\theta})\right| &\le Ch^2 \|\dot{\bar{\boldsymbol{U}}}\|_{H^1(\Gamma_h)} \|\boldsymbol{\theta}\|_{H^1(\Gamma_h)} \le Ch^2\left(\|\dot{\boldsymbol{\rho}}\|_{H^1(\Gamma_h)} + \left\|\dot{\boldsymbol{u}}^{-\ell}\right\|_{H^1(\Gamma_h)}\right)\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)}\\ &\le Ch^2\left(C(\boldsymbol{u})h + C\|\dot{\boldsymbol{u}}\|_{H^1(\Gamma)}\right)\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)} \le C(\boldsymbol{u})h^2 \|\boldsymbol{\theta}\|_{H^1(\Gamma_h)}.\end{aligned} \tag{55}$$

The Cauchy–Schwarz inequality, (29), the geometric estimate (33) and the stability bound (36) yield

$$\left|\int_\Gamma \left(1-\frac{1}{\delta_h^\ell}\right)\dot{\boldsymbol{u}}:\boldsymbol{\theta}^\ell\right| \le \left\|1-\frac{1}{\delta_h^\ell}\right\|_{L^\infty(\Gamma)} \|\dot{\boldsymbol{u}}\|_{L^2(\Gamma)} \|\boldsymbol{\theta}\|_{L^2(\Gamma_h)} \le C(\boldsymbol{u})h^2 \|\boldsymbol{\theta}\|_{L^2(\Gamma_h)}. \tag{56}$$

Combining (49)– (56), using (27), (29), (30) and (35), we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\boldsymbol{\theta}\|_h^2 + m|\nabla_{\Gamma_h}\boldsymbol{\theta}\|_{L^2(\Gamma_h)}^2 \le C(\boldsymbol{u})\left(h^2+\|\boldsymbol{\theta}\|_h\right)\|\boldsymbol{\theta}\|_{H^1(\Gamma_h)} \le C(\boldsymbol{u},m)\left(h^4+|\boldsymbol{\theta}\|_h^2\right) + m|\boldsymbol{\theta}\|_{H^1(\Gamma_h)}^2,$$

where $m = \min(\mathrm{eig}(\boldsymbol{D}^s))$. Cancelling $m|\nabla_{\Gamma_h}\boldsymbol{\theta}\|_{L^2(\Gamma_h)}^2$ on both sides and using (27), we have that $\frac{\mathrm{d}}{\mathrm{d}t}|\boldsymbol{\theta}\|_h^2 \le C(\boldsymbol{u})h^4 + C(\boldsymbol{u})|\boldsymbol{\theta}\|_h^2$. Using Grönwall's lemma, the assumption $\|\boldsymbol{\theta}_0^\ell\|_{L^2(\Gamma)} \le Ch^2$, (27) and (29), we obtain $|\boldsymbol{\theta}^\ell\|_{L^2(\Gamma)}^2 \le C(\boldsymbol{u},T)h^4$, which yields the desired result. □

Similarly to the approach employed in [11] and [32], one obtains the following $L^\infty([0,T],L^2(\Gamma))$ error estimate for the fully-discrete solution (10) as follows.

Theorem 5 (*Error estimate for the fully-discrete solution (10)*). *Assume that Σ is a bounded invariant region for* (3) and (10), *$\boldsymbol{f} \in C^2(\Sigma)$ and $\boldsymbol{u}_0, \boldsymbol{U}_0 \in \Sigma$. If the solution $\boldsymbol{u}$ of* (3) *and its time derivative $\dot{\boldsymbol{u}}$ are $L^\infty([0,T];H^2(\Gamma))$, $\ddot{\boldsymbol{u}}$ is $L^\infty([0,T];L^2(\Gamma))$ and $\|\boldsymbol{u}_0 - \boldsymbol{U}_0^\ell\|_{L^2(\Gamma)} \le ch^2$, then the following estimate holds*

$$\|\boldsymbol{u}^n - \boldsymbol{U}^{\ell,n}\|_{L^2(\Gamma)} \le C(\boldsymbol{u},T)(h^2+\tau), \qquad n = 0,\ldots,N_T,$$

where $\boldsymbol{u}^n$ is the exact solution at time $t_n := n\tau$ and $C(\boldsymbol{u},T)$ is a constant depending on $\boldsymbol{u}$ and T.

Proof. Following [17],Theorem 8 let us write the error as $\boldsymbol{U}^{\ell,n} - \boldsymbol{u}^n = (\boldsymbol{U}^{\ell,n} - \bar{\boldsymbol{U}}^{\ell,n}) + (\bar{\boldsymbol{U}}^{\ell,n} - \boldsymbol{u}^n) =: \boldsymbol{\theta}^{\ell,n} + \boldsymbol{\rho}^{\ell,n}$ and the discrete time derivative of any function $\boldsymbol{\phi}:\Gamma_h\times[0,T]\to\mathbb{R}^r$ as $\bar{\partial}\boldsymbol{\phi}^n := \frac{\boldsymbol{\phi}^n-\boldsymbol{\phi}^{n-1}}{\tau}$. Since $\boldsymbol{u}$ and $\dot{\boldsymbol{u}}$ are $L^\infty([0,T],H^2(\Gamma))$, from (29), (30), (44) and (45), we have that

$$\|\boldsymbol{\rho}^n\|_{L^2(\Gamma_h)} \le C\|\boldsymbol{\rho}^{\ell,n}\|_{L^2(\Gamma)} = \|\bar{\boldsymbol{U}}^{\ell,n} - \boldsymbol{u}^n\|_{L^2(\Gamma)} \le ch^2\left\|\boldsymbol{u}^n\right\|_{H^2(\Gamma)}, \tag{57}$$

$$\|\dot{\boldsymbol{\rho}}^n\|_{L^2(\Gamma_h)} + h\|\nabla_{\Gamma_h}\dot{\boldsymbol{\rho}}^n\|_{L^2(\Gamma_h)} \le ch^2(\|\boldsymbol{u}^n\|_{H^2(\Gamma)} + \|\dot{\boldsymbol{u}}^n\|_{H^2(\Gamma)}). \tag{58}$$

It remains to show the convergence for $\boldsymbol{\theta}^{\ell,n}$. To this end, we derive an estimate for $\boldsymbol{\theta}^n$ in the $L^2(\Gamma_h)$ norm and then use (27) and (29) to estimate $\|\boldsymbol{\theta}^{\ell,n}\|_{L^2(\Gamma)}$. The continuous problem (3) and the fully-discrete formulation (10), the definition of Ritz projection (43), and the relation (32), imply that

$$\begin{aligned}&\int_{\Gamma_h} I_h(\bar{\partial}\boldsymbol{\theta}^n:\boldsymbol{\varphi}^n) + \int_{\Gamma_h}\boldsymbol{D}\nabla_{\Gamma_h}\boldsymbol{\theta}^n:\nabla_{\Gamma_h}\boldsymbol{\varphi}^n = -\varepsilon_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell,n-1}),\boldsymbol{\varphi}^n) - \int_\Gamma\left(1-\frac{1}{\delta_h^\ell}\right)\boldsymbol{f}(\boldsymbol{u}^{n-1}):\boldsymbol{\varphi}^{\ell,n}\\ &\quad+\int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{U}^{n-1})-\boldsymbol{f}(\boldsymbol{u}^{-\ell,n-1})):\boldsymbol{\varphi}^n) + \int_\Gamma(\boldsymbol{f}(\boldsymbol{u}^{n-1})-\boldsymbol{f}(\boldsymbol{u}^n)):\boldsymbol{\varphi}^{\ell,n}\\ &\quad-\int_{\Gamma_h}\bar{\partial}\boldsymbol{\rho}^n:\boldsymbol{\varphi}^n + \varepsilon_h(\bar{\partial}\bar{\boldsymbol{U}}^n,\boldsymbol{\varphi}^n) - \int_{\Gamma_h}(\bar{\partial}-\partial_t)\boldsymbol{u}^{-\ell,n}:\boldsymbol{\varphi}^n + \int_\Gamma\left(1-\frac{1}{\delta_h^\ell}\right)\dot{\boldsymbol{u}}^n:\boldsymbol{\varphi}^{\ell,n}.\end{aligned} \tag{59}$$

In (59) we choose $\boldsymbol{\varphi}^n = \boldsymbol{\theta}^n$. For the first term in (59) we observe that, from Young's inequality,

$$\int_{\Gamma_h} I_h\left(\bar{\partial}\boldsymbol{\theta}^n:\boldsymbol{\theta}^n\right) \ge \frac{1}{2\tau}(|\boldsymbol{\theta}^n\|_h^2 - |\boldsymbol{\theta}^{n-1}\|_h^2). \tag{60}$$

We estimate the single terms on the right-hand side of (59) in turn. From the Cauchy–Schwarz inequality, the Lipschitz continuity of $\boldsymbol{f}$, the definition of $\boldsymbol{\theta}^n$, (27) and (57), it follows that

$$\begin{aligned}&\left|\int_{\Gamma_h} I_h((\boldsymbol{f}(\boldsymbol{U}^{n-1})-\boldsymbol{f}(\boldsymbol{u}^{-\ell,n-1})):\boldsymbol{\theta}^n)\right| \le C\|\boldsymbol{U}^{n-1}-\boldsymbol{u}^{-\ell,n-1}\|_h\|\boldsymbol{\theta}^n\|_h\\&\quad\le C(\|\boldsymbol{\rho}^{n-1}\|_{L^2(\Gamma)}+\|\boldsymbol{\theta}^{n-1}\|_h)\|\boldsymbol{\theta}^n\|_h \le C(\boldsymbol{u})(h^2+\|\boldsymbol{\theta}^{n-1}\|_h)\|\boldsymbol{\theta}^n\|_h.\end{aligned}\tag{61}$$

From the estimate (28) for ε_h and (31), we obtain that

$$\begin{aligned}&\left|\varepsilon_h(\boldsymbol{f}(\boldsymbol{u}^{-\ell,n-1}),\boldsymbol{\theta}^n)\right| \le Ch^2\|\boldsymbol{f}(\boldsymbol{u}^{-\ell,n-1})\|_{H_h^2(\Gamma_h)}\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)}\\&\quad\le Ch^2\|\boldsymbol{f}(\boldsymbol{u}^{n-1})\|_{H^2(\Gamma)}\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)} \le Ch^2\|\boldsymbol{f}\|_{\mathcal{C}^2(\Sigma)}\|\boldsymbol{u}^{n-1}\|_{H^2(\Gamma)}\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)} \le Ch^2\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)},\end{aligned}\tag{62}$$

where we have exploited the regularity assumptions $\boldsymbol{f}\in\mathcal{C}^2(\Sigma)$ and $\boldsymbol{u}\in L^\infty([0,T],H^2(\Gamma))$. Since $\boldsymbol{f}$ is Lipschitz over the compact region Σ then $\boldsymbol{f}\in L^\infty(\Sigma)$. This fact, together with Cauchy–Schwarz inequality, (29) and the geometric estimate (33), yields

$$\left|\int_\Gamma\left(1-\frac{1}{\delta_h^\ell}\right)\boldsymbol{f}(\boldsymbol{u}^{n-1}):\boldsymbol{\theta}^{\ell,n}\right| \le \left\|1-\frac{1}{\delta_h^\ell}\right\|_{L^\infty(\Gamma)}\left\|\boldsymbol{f}(\boldsymbol{u}^{n-1})\right\|_{L^2(\Gamma_h)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma)} \le Ch^2\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}.\tag{63}$$

The Cauchy–Schwarz inequality yields, together with (29) and the stability estimate (37),

$$\begin{aligned}&\left|\int_\Gamma(\boldsymbol{f}(\boldsymbol{u}^{n-1})-\boldsymbol{f}(\boldsymbol{u}^n)):\boldsymbol{\theta}^{\ell,n}\right| \le C\|\boldsymbol{u}^n-\boldsymbol{u}^{n-1}\|_{L^2(\Gamma)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)} = \left\|\int_{t_{n-1}}^{t_n}\dot{\boldsymbol{u}}\right\|_{L^2(\Gamma)}\|\boldsymbol{\theta}^n\|_{L^2(\Gamma_h)}\\&\quad\le\|\boldsymbol{\theta}^n\|_{L^2(\Gamma_h)}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{u}}\|_{L^2(\Gamma)} \le \tau\,\|\dot{\boldsymbol{u}}\|_{L^\infty([0,T],L^2(\Gamma))}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)} = C(\boldsymbol{u})\tau\|\boldsymbol{\theta}^n\|_{L^2(\Gamma_h)}.\end{aligned}\tag{64}$$

The Cauchy–Schwarz inequality and the estimate (58) for $\dot{\boldsymbol{\rho}}$ yield

$$\begin{aligned}&\left|\int_{\Gamma_h}\bar{\partial}\boldsymbol{\rho}^n:\boldsymbol{\theta}^n\right| \le C\left\|\bar{\partial}\boldsymbol{\rho}^n\right\|_{L^2(\Gamma_h)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)} = \frac{C}{\tau}\left\|\int_{t_{n-1}}^{t_n}\dot{\boldsymbol{\rho}}\right\|_{L^2(\Gamma_h)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}\\&\quad\le\frac{C}{\tau}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\rho}}\|_{L^2(\Gamma_h)} \le C\left\|\dot{\boldsymbol{\rho}}\right\|_{L^\infty([0,T],L^2(\Gamma_h))}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)} \le C(\boldsymbol{u})h^2\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}.\end{aligned}\tag{65}$$

From the estimate (28) for ε_h, the estimate (58) for $\dot{\boldsymbol{\rho}}$, the equivalences (29) and (30), the triangle inequality and $\bar{\partial}\bar{\boldsymbol{U}}^n,\boldsymbol{\theta}^n\in S_h$ we obtain

$$\begin{aligned}&\left|\varepsilon_h(\bar{\partial}\bar{\boldsymbol{U}}^n,\boldsymbol{\theta}^n)\right| \le Ch^2\|\bar{\partial}\bar{\boldsymbol{U}}^n\|_{H^1(\Gamma_h)}\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)} \le \frac{Ch^2}{\tau}\left\|\boldsymbol{\theta}^n\right\|_{H^1(\Gamma_h)}\int_{t_{n-1}}^{t_n}\|\dot{\bar{\boldsymbol{U}}}\|_{H^1(\Gamma_h)}\\&\quad\le Ch^2\|\dot{\bar{\boldsymbol{U}}}\|_{L^\infty([0,T],H^1(\Gamma_h))}\left\|\boldsymbol{\theta}^n\right\|_{H^1(\Gamma_h)} \le \left(\|\dot{\boldsymbol{\rho}}\|_{L^\infty([0,T],H^1(\Gamma_h))}\|\dot{\boldsymbol{u}}^{-\ell}\|_{L^\infty([0,T],H^1(\Gamma_h))}\right)\left\|\boldsymbol{\theta}^n\right\|_{H^1(\Gamma_h)}\\&\quad\le Ch^2\left(C(\boldsymbol{u})h+C\left\|\dot{\boldsymbol{u}}\right\|_{L^\infty([0,T],H^1(\Gamma))}\right)\left\|\boldsymbol{\theta}^n\right\|_{H^1(\Gamma_h)} \le C(\boldsymbol{u})h^2\left\|\boldsymbol{\theta}^n\right\|_{H^1(\Gamma_h)}.\end{aligned}\tag{66}$$

The Cauchy–Schwarz inequality and (29) give rise to the following inequalities

$$\begin{aligned}&\left|\int_{\Gamma_h}(\bar{\partial}-\partial_t)\boldsymbol{u}^{-\ell,n}:\boldsymbol{\theta}^n\right|\le C\left\|(\bar{\partial}-\partial_t)\boldsymbol{u}^n\right\|_{L^2(\Gamma)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}\le\frac{C}{\tau}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{u}}(t)-\dot{\boldsymbol{u}}(t_n)\|_{L^2(\Gamma)}\,\mathrm{d}t\\&\quad\le\frac{C}{\tau}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}\int_{t_{n-1}}^{t_n}\int_t^{t_n}\|\ddot{\boldsymbol{u}}(s)\|\,\mathrm{d}s\mathrm{d}t \le C\tau\left\|\ddot{\boldsymbol{u}}\right\|_{L^\infty([0,T],L^2(\Gamma))}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}=C(\boldsymbol{u})\tau\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)},\end{aligned}\tag{67}$$

where we have exploited the assumption that $\ddot{\boldsymbol{u}}\in L^\infty([0,T],L^2(\Gamma))$. The Cauchy–Schwarz inequality, (29), the geometric estimate (33) and the stability bound (36) yield

$$\left|\int_\Gamma\left(1-\frac{1}{\delta_h^\ell}\right)\dot{\boldsymbol{u}}^n:\boldsymbol{\theta}^{\ell,n}\right| \le \left\|1-\frac{1}{\delta_h^\ell}\right\|_{L^\infty(\Gamma)}\left\|\dot{\boldsymbol{u}}^n\right\|_{L^2(\Gamma)}\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)} \le C(\boldsymbol{u})h^2\left\|\boldsymbol{\theta}^n\right\|_{L^2(\Gamma_h)}.\tag{68}$$

Combining (59)– (68), using (27) and (35) and Young's inequality we get

$$\begin{aligned}&\frac{1}{2\tau}\left(|\boldsymbol{\theta}^n\|_h^2-|\boldsymbol{\theta}^{n-1}\|_h^2\right)+m|\nabla_{\Gamma_h}\boldsymbol{\theta}^n\|_{L^2(\Gamma_h)}^2 \le C(\boldsymbol{u})\left(h^2+\tau+\|\boldsymbol{\theta}^{n-1}\|_h\right)\|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)}\\&\quad\le C(\boldsymbol{u},m)\left(h^4+\tau^2+|\boldsymbol{\theta}^{n-1}\|_h^2\right)+m|\boldsymbol{\theta}^n\|_{H^1(\Gamma_h)}^2,\end{aligned}\tag{69}$$

where $m=\min(\mathrm{eig}(\boldsymbol{D}^s))$, from which, cancelling $\|\nabla_{\Gamma_h}\boldsymbol{\theta}^n\|_{L^2(\Gamma_h)}$ on both sides of (69), and using (27), we have that

$$|\boldsymbol{\theta}^n\|_h^2 \le (1+C(\boldsymbol{u})\tau)|\boldsymbol{\theta}^{n-1}\|_h^2+C(\boldsymbol{u})\tau(h^4+\tau^2).\tag{70}$$

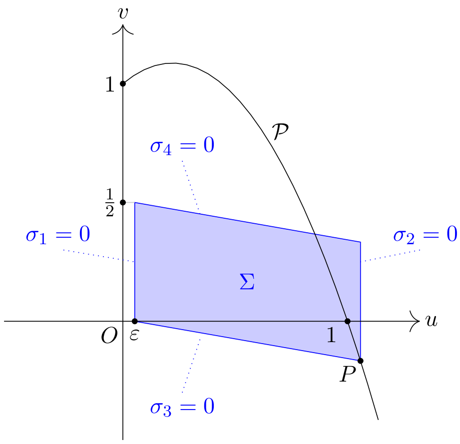


Fig. 1. *Test* 1: invariant parallelogram Σ for the Rosenzweig–MacArthur system (71) with parameters as stated in the text and $\varepsilon = $ 1e−7. The edges are represented as the zero-level sets of the functions $\sigma_i, i = 1, \dots, 4$ defined in the text. The slope of the slanting edges is $-\frac{1}{6}$. The corner P lies on the nullcline $\mathcal{P}$ of the kinetic for u.

By recursively applying (70), taking into account the assumption that $\|\boldsymbol{\theta}^0\|_{L^2(\Gamma)} \leq Ch^2$, and then using (27) and (29), we obtain $|\boldsymbol{\theta}^{\ell,n}\|^2_{L^2(\Gamma)} \leq C(\boldsymbol{u})(h^4 + \tau^2)$, which yields the desired result. □

In summary, the previous theorems entail that our semi- and fully-discrete schemes exhibit optimal convergence rates that are quadratic in the mesh size and linear in the timestep.

4. Numerical tests

In this section we present two examples to show that the LSFEM–IMEX Euler full discretisation of RCD systems (i) fulfils the conditions given in Theorem 2 for the existence of invariant polytopes, whilst the SFEM–IMEX Euler full discretisation does not (Test 1) and (ii) can be applied for the approximation of Turing patterns on surfaces, in good agreement with the results obtained with another method in [33] (Test 2). The simulations have been carried out using MATLAB. The linear system arising at each timestep is solved with MATLAB's "backslash" command. The code is available on request.

4.1. Test 1: Invariant parallelogram

In this experiment we consider the RCD system with non-dimensional Rosenzweig–MacArthur kinetics [25,34] and linear cross-diffusion given by

$$\begin{cases} u_t - d_{uu}\Delta_\Gamma u - d_{uv}\Delta_\Gamma v = au(1-u) - b\dfrac{uv}{u+\alpha}, \\ v_t - d_{vu}\Delta_\Gamma u - d_{vv}\Delta_\Gamma v = c\dfrac{uv}{u+\alpha} - dv, \end{cases} \tag{71}$$

on the unit sphere Γ, where α, a, b, c and d are positive constants.

In the absence of cross-diffusion, this model has been solved in [25] on a planar domain. To the best of the authors' knowledge, until now there is no discussion about the existence of an invariant region at the discrete level. In the present example we show that the IMEX–LSFEM full discretisation of system (71) possesses an invariant parallelogram in the presence of linear cross-diffusion with no modifications of the kinetics. For the reaction kinetics, we choose the following parameters α = 1e-3, a = 10, b = 1e-2, c = 1, d = 2.2. For the diffusion coefficients, we choose $\begin{pmatrix} d_{uu} & d_{uv} \\ d_{vu} & d_{vv} \end{pmatrix} = \begin{pmatrix} 6e\text{-}2 & 0 \\ 1e\text{-}2 & 1.2e\text{-}1 \end{pmatrix}$. It is possible to verify that the parallelogram Σ defined by

$$\Sigma = \{(u, v) \in \mathbb{R}^2 \mid \sigma_l(u, v) \geq 0,\ l = 1, \dots, 4\}, \tag{72}$$

where the affine functions $\sigma_l, l = 1, \dots, 4$ are given by $\sigma_1(u, v) = u - \varepsilon$, $\sigma_2(u, v) = 6 - 5\alpha + \sqrt{(6-5\alpha)^2 + 24\alpha(6-\varepsilon)} - 12u$, $\sigma_3(u, v) = u + 6v - \varepsilon$, $\sigma_4(u, v) = 3 + \varepsilon - u - 6v$, with ε = 1e-7, is an invariant region for system (71). Σ is depicted in Fig. 1.

Table 1
Test 1: Invariance analysis for the SFEM solution. The solution blows up on the first five meshes. On the three finest meshes the numerical solution stays bounded, though still violating the invariant parallelogram Σ in Fig. 1.

i	N	h	$\min_{\Gamma_h\times[\tau,5]}\sigma_1(U)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_2(U)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_3(U,V)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_4(U,V)$
0	126	4.013e−01	−7.503e+271	−4.733e+269	−7.460e+271	−7.730e+268
1	258	2.863e−01	−1.538e+305	−1.935e+303	−1.521e+305	−3.086e+302
2	516	2.026e−01	−1.871e−02	2.198e−02	−4.275e+00	−1.707e+00
3	1062	1.414e−01	−1.704e−02	−1.901e−01	−2.756e−01	−1.588e+00
4	2094	1.007e−01	−1.424e−02	2.198e−02	−8.777e+00	1.613e−02
5	4242	7.082e−02	−1.288e−02	2.198e−02	1.553e−01	−5.531e+00
6	8370	5.041e−02	−9.164e−03	2.198e−02	1.553e−01	−2.390e+00
7	16962	3.542e−02	−6.391e−03	2.198e−02	1.553e−01	7.946e−03

Table 2
Test 1: Invariance analysis for the LSFEM solution. The solution stays in the invariant parallelogram Σ in (72) on all considered meshes.

i	N	h	$\min_{\Gamma_h\times[\tau,5]}\sigma_1(U)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_2(U)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_3(U,V)$	$\min_{\Gamma_h\times[\tau,5]}\sigma_4(U,V)$
0	126	4.013e−01	6.667e−10	2.198e−02	1.619e−01	2.033e−02
1	258	2.863e−01	6.667e−10	2.198e−02	1.568e−01	7.509e−02
2	516	2.026e−01	6.667e−10	2.198e−02	1.556e−01	1.158e−01
3	1062	1.414e−01	6.667e−10	2.198e−02	1.553e−01	8.724e−03
4	2094	1.007e−01	6.667e−10	2.198e−02	1.553e−01	1.496e−02
5	4242	7.082e−02	6.667e−10	2.198e−02	1.553e−01	8.220e−03
6	8370	5.041e−02	6.667e−10	2.198e−02	1.553e−01	1.326e−02
7	16962	3.542e−02	6.667e−10	2.198e−02	1.553e−01	8.201e−03

The invariance of Σ means that σ_l, $l = 1, \dots, 4$, defined above, are positive for all times after discretisation. The $H^1(\Gamma)$ initial datum

$$u_0(x, y, z) = \begin{cases} \varepsilon + (1-\varepsilon)\sqrt{1 - \dfrac{x^2+y^2}{r^2}} & \text{if } x^2+y^2 \le r^2,\ z > 0, \\ \varepsilon & \text{elsewhere}, \end{cases} \tag{73}$$

$$v_0(x, y, z) = \frac{a\alpha}{3b}, \qquad \forall (x, y, z) \in \Gamma, \tag{74}$$

with $r = 0.2$, is contained in the invariant region Σ. It is easy to verify that, on Σ, the Lipschitz constants L_1 and L_2 of the kinetics in (71) satisfy

$$L_1 < \tilde{L}_1 := \sqrt{2}\left(3a + \frac{b}{2\alpha}\right) \approx 49.4975, \text{ and } L_2 < \tilde{L}_2 := \sqrt{2}\left(\frac{c}{2\alpha} + \frac{d}{2}\right) \approx 708.6624.$$

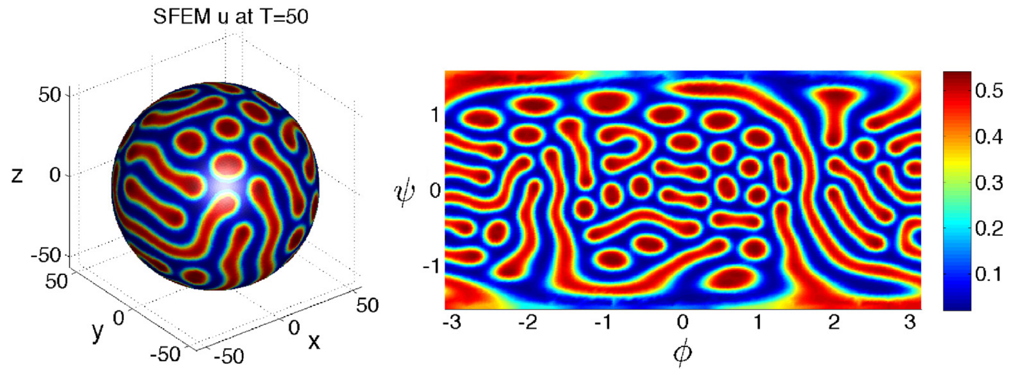
The stability condition (21) on the timestep is fulfilled if we choose

$$\tau \le \bar{\tau} := \frac{1}{\max\left(\tilde{L}_1, \sqrt{\frac{1}{37}\tilde{L}_1^2 + \frac{36}{37}\tilde{L}_2^2}\right)} \approx 1.43e\text{-}3. \tag{75}$$

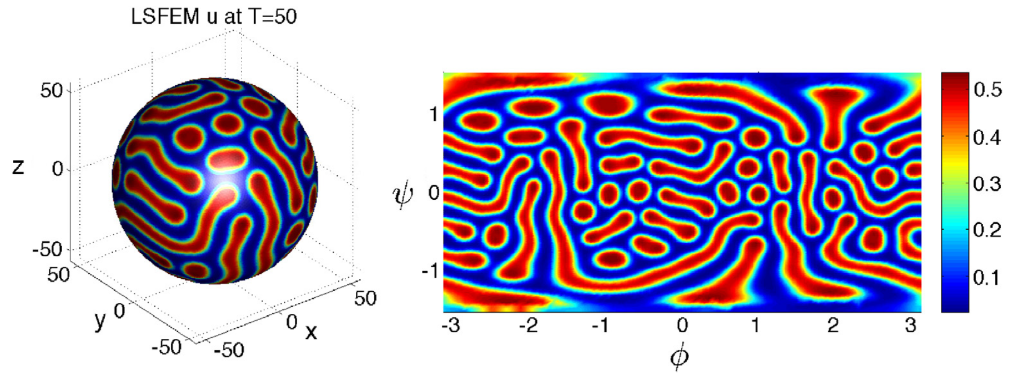
We solve the problem with a fixed timestep $\tau = 1e\text{-}3$ until the final time $T = 5$, on a sequence of eight meshes Γ_i, $i = 0, \dots, 7$ with decreasing meshsizes $h_i \approx \frac{h_0}{(\sqrt{2})^i}$, $h_0 = 0.4013$, so that, for all $i = 0, \dots, 6$, the number of nodal points of Γ_{i+1} is approximately double that of Γ_i. For all $i = 0, \dots, 7$, the minima of σ_l, $l = 1, \dots, 4$, defined above are shown in Table 1 for SFEM and in Table 2 for LSFEM. We observe that the LSFEM solution is in Σ at all times, whilst the SFEM solution without lumping escapes Σ on all considered meshes. Furthermore, the SFEM exhibits a stability threshold: the numerical solution blows up on meshes Γ_i, $i = 0, 1$, while it appears to stay bounded on the finer meshes Γ_i, $i = 2, \dots, 7$. It is worth noting that the timestep restriction (75) is only a sufficient condition for the IMEX–LSFEM scheme to possess an invariant region. In fact, we have carried out the above invariance test with larger timesteps and we have observed that the IMEX–LSFEM admits Σ as an invariant region on all meshes Γ_i also for larger values of τ, that is $1e\text{-}3 \le \tau \le 0.1$, while for $\tau = 0.2$ the method violates Σ on all meshes Γ_i.

4.2. Test 2: Pattern formation

In this experiment, we solve the RCDS with Rosenzweig–MacArthur kinetics in (71) with diffusion coefficients, parameters and final time given by, respectively, $\begin{pmatrix} d_{uu} & d_{uv} \\ d_{vu} & d_{vv} \end{pmatrix} = \begin{pmatrix} 100 & 100 \\ 400 & 500 \end{pmatrix}$, $\alpha = \frac{11}{15}$, $a = 1$, $b = \frac{2}{3}$, $c = \frac{2}{30}$, $d = \frac{11}{1000}$, and $T = 50$. This choice is equivalent, by rescaling time, to the parameter choice in Fig. 4A of [33] and leads to Turing instability, as proven therein. The initial condition is a spatially random perturbation, of amplitude $1e\text{-}5$, of the homogeneous steady state $(u^*, v^*) := \left(\frac{d\alpha}{c-d}, \frac{a}{b}(1 - \frac{d\alpha}{c-d})(\alpha + \frac{d\alpha}{c-d})\right) = (0.1935, 1.0406)$. Since, in [33], the problem is solved on the square $[0, 200]^2$, we



(a) SFEM



(b) LSFEM

Fig. 2. *Test 2*: u-component of the Rosenzweig–MacArthur system (71) on the sphere with parameters as stated in the text. The sphere of radius $R = 100/\sqrt{\pi}$ is approximated with a triangular mesh of $N = 16962$ gridpoints, the timestep used is $\tau = 1\text{e}{-2}$. The planar deployments of the numerical solutions, through spherical coordinates (ϕ, ψ), are shown on the right side of each panel.

consider a sphere of the same area, thus with radius $R = \frac{100}{\sqrt{\pi}}$. We solve the system with SFEM and LSFEM on a mesh with $N = 16962$ gridpoints and timestep $\tau = 1e$-2. The solutions at the final time $T = 50$ are shown in Fig. 2(a) for SFEM and in Fig. 2(b) for LSFEM, respectively. We observe that (i) starting from the same initial datum, SFEM and LSFEM exhibit almost the same final pattern and (ii) with SFEM and LSFEM, we obtain the same kind of patterns obtained in [33] by using finite differences in space (on the planar domain).

5. Conclusions

In this study we have considered a lumped surface finite element method (LSFEM) for systems of arbitrarily many semilinear parabolic equations with linear cross-diffusion on stationary surfaces, by extending its counterpart without cross-diffusion studied in [17]. Time discretisation is carried out by applying the Euler IMEX scheme in time that approximates all diffusion terms implicitly. In Theorem 1 we have shown that, under the assumption of Delaunay regularity for the mesh, provided the diffusion coefficients are compatible with the orientation of the hyper-faces of the polytope, the strictly inward flux condition (13) and the compatibility condition (14) are sufficient for a polytope in the phase space to be invariant for the spatially discrete scheme. For the fully-discrete problem arising from Euler IMEX scheme we have shown in Theorem 2 that, under the timestep restriction (21) involving the Lipschitz constants of the reaction kinetics, conditions (13)-(14) are still sufficient to ensure a hyper-rectangle to be invariant. To the best of the authors' knowledge, Theorems 1 and 2 are a novelty even on planar domains.

For both the semi- and fully-discrete formulations of the RCDSs considered in Section 2, an optimal $L^2(\Gamma)$ error bound has been proven in Theorems 4 and 5 in Section 3. The numerical tests in Section 4 confirm our theoretical findings. The usefulness of LSFEM is illustrated in *Tests 1* and *2*. In particular, we have shown that in the absence of lumping, the numerical solution of a classical predator–prey model with the addition of cross-diffusion blows-up instead of being bounded in the invariant parallelogram.

Emerging applications encourage the extension of the present study to the case of *evolving* surfaces. Another extension, motivated by a number of existing models in the literature, is towards the class of systems with *nonlinear* cross-diffusion.

This latter extension is challenging in that (i) nonlinear diffusion needs a different discretisation and (ii) the monotonicity properties of such systems are not as well-understood as in the case of linear diffusion. For these reasons, we believe that a different numerical analysis is needed for systems with nonlinear diffusion. The two aforementioned extensions are beyond the scope of this work and will be addressed in future studies.

Acknowledgements

All the authors (AM, IS, CV, MF) thank the Isaac Newton Institute for Mathematical Sciences for its hospitality during the programme (Coupling Geometric PDEs with Physics for Cell Morphology, Motility and Pattern Formation; EPSRC EP/K032208/1). This work (AM) has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 642866. AM and CV acknowledge support from the Engineering and Physical Sciences Research Council (EP/J016780/1) on *Modelling, analysis and simulation of spatial patterning on evolving biological surfaces* and the Leverhulme Trust Research Project Grant (RPG-2014-149) on *Unravelling new mathematics for 3D cell migration*. AM was partially supported by a fellowship from the Simons Foundation. AM is a Royal Society Wolfson Research Merit Award Holder, generously funded by the Wolfson Foundation. MF, IS, and CV would like to thank the Department of Mathematics of the University of Sussex for its hospitality during the preparation of the present work.

Appendix

Lemma A (*Preservation of linear constraints*). *Given $r, N \in \mathbb{N}$, let $\boldsymbol{B} \in \mathbb{R}^{N,N}$ be a matrix with real nonnegative eigenvalues such that $b_{ij} \le 0$ for $i \neq j$, let $\boldsymbol{H} \in \mathbb{R}^{r,r}$ be a (possibly non-symmetric) positive definite matrix, let $\boldsymbol{n}$ be a left eigenvector of $\boldsymbol{H}$ with real eigenvalue λ, let $\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_r \in \mathbb{R}^N$ and let $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_r)^T \in \mathbb{R}^{rN}$. If $(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\boldsymbol{\zeta} \ge \boldsymbol{0}$, then it holds that $(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)(\boldsymbol{I}_{rN} + \boldsymbol{H} \otimes \boldsymbol{B})^{-1}\boldsymbol{\zeta} \ge \boldsymbol{0}$.*

Proof. For $\mu > 0$, we consider the matrix $K := \mu \boldsymbol{I}_{rN} - (\boldsymbol{I}_{rN} + \boldsymbol{H} \otimes \boldsymbol{B}) = (\mu - 1)\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B}$. Now, $\bar{\lambda}$ is an eigenvalue of $\boldsymbol{H} \otimes \boldsymbol{B}$ if and only if $\bar{\bar{\lambda}} := \mu - 1 - \bar{\lambda}$ is an eigenvalue of K, in fact $\det(\bar{\lambda}\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B}) = \det((\mu - 1 - \bar{\lambda})\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B}) = \boldsymbol{0}$. Notice that

$$\mathrm{Re}(\bar{\bar{\lambda}}) = \mu - 1 - \mathrm{Re}(\bar{\lambda}), \qquad \mathrm{Im}(\bar{\bar{\lambda}}) = -\mathrm{Im}(\bar{\lambda}). \tag{76}$$

Since, from the positive definiteness, the eigenvalues of $\boldsymbol{H}$ have positive real part and, by assumption, $\boldsymbol{B}$ has real nonnegative eigenvalues, then the eigenvalues of $\boldsymbol{H} \otimes \boldsymbol{B}$ have nonnegative real part, namely $\mathrm{Re}(\bar{\lambda}) \ge 0$. Consequently, (76) implies $|\mathrm{Re}(\bar{\bar{\lambda}})| \le \mu - 1$, which, in combination with (76), yields $|\bar{\bar{\lambda}}|^2 \le (\mu - 1)^2 + \left(\mathrm{Im}(\bar{\lambda})\right)^2$. It follows that, by choosing $\mu > \max_{\bar{\lambda} \in \mathrm{eig}(\boldsymbol{H} \otimes \boldsymbol{B})} \frac{(\mathrm{Im}(\bar{\lambda}))^2 + 1}{2}$, the spectral radius of K is less than μ, then the spectral radius of $\mu^{-1}K$ is less than 1, and thus $\mu^{-1}K$ may be expressed as the sum of a geometric series. Assume now $(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\boldsymbol{\zeta} \ge \boldsymbol{0}$. Then

$$\begin{aligned}
(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)(\boldsymbol{I}_{rN} + \boldsymbol{H} \otimes \boldsymbol{B})^{-1}\boldsymbol{\zeta} &= (\boldsymbol{n}^T \otimes \boldsymbol{I}_N)(\mu \boldsymbol{I}_{rN} - K)^{-1}\boldsymbol{\zeta} = (\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\mu^{-1}(\boldsymbol{I}_{rN} - \mu^{-1}K)^{-1}\boldsymbol{\zeta} \\
&= (\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\mu^{-1}\sum_{j=0}^{+\infty} \mu^{-j}K^j\boldsymbol{\zeta} = \sum_{j=0}^{+\infty} \mu^{-1-j}(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)((\mu - 1)\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B})^j\boldsymbol{\zeta}.
\end{aligned}$$

We need to prove that, for all $j \in \mathbb{N}$, $(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)((\mu - 1)\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B})^j\boldsymbol{\zeta} \ge \boldsymbol{0}$. However, by induction, it suffices to prove that

$$(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)((\mu - 1)\boldsymbol{I}_{rN} - \boldsymbol{H} \otimes \boldsymbol{B})\boldsymbol{\zeta} \ge \boldsymbol{0}. \tag{77}$$

From the properties of the Kronecker product, the left-hand side in (77) can be rearranged as

$$[(\mu - 1)\boldsymbol{n}^T \otimes \boldsymbol{I}_N - (\boldsymbol{n}^T\boldsymbol{H}) \otimes (\boldsymbol{I}_N\boldsymbol{B})]\boldsymbol{\zeta} = [(\mu - 1)\boldsymbol{n}^T \otimes \boldsymbol{I}_N - \lambda\boldsymbol{n}^T \otimes \boldsymbol{B}]\boldsymbol{\zeta}. \tag{78}$$

Claim (77) can now be written componentwise as

$$(\mu - 1)((\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\boldsymbol{\zeta})_i \ge \lambda((\boldsymbol{n}^T \otimes \boldsymbol{B})\boldsymbol{\zeta})_i, \quad i = 1, \dots, N. \tag{79}$$

We recast the left-hand side of (79) as

$$(\mu - 1)\sum_{k=1}^{r} n_k\zeta_{k,i}, \tag{80}$$

and the right-hand side of (79) as

$$\lambda\sum_{k=1}^{r} n_k(\boldsymbol{B}\boldsymbol{\zeta}_k)_i = \lambda\sum_{k=1}^{r} n_k\sum_{j=1}^{N} b_{ij}\zeta_{k,j} = \lambda\sum_{j=1}^{N} b_{ij}\sum_{k=1}^{r} n_k\zeta_{k,j} \le \lambda b_{ii}\sum_{k=1}^{r} n_k\zeta_{k,i}, \tag{81}$$

where, in the inequality, we have exploited the assumption that $b_{ij} \le 0$ for $i \neq j$ and $(\boldsymbol{n}^T \otimes \boldsymbol{I}_N)\boldsymbol{\zeta} \ge \boldsymbol{0}$. Now it suffices to prove that the right-hand side of (81) is less than or equal to (80) for all $i = 1, \dots, N$, which is true by enforcing $\mu \ge \lambda\max_{i=1,\dots,N}(b_{ii}) + 1$. □

Lemma B (*Zero discrete diffusion of spatially uniform states*)**.** *Let* $\overline{\boldsymbol{M}}$, $\boldsymbol{A}$, $\boldsymbol{D}$ *be the lumped mass matrix, the stiffness matrix and the diffusivity matrix introduced above, let* $\tau > 0$ *and let* $\boldsymbol{v} \in \mathbb{R}^r$ *be a column vector. Then* $(\boldsymbol{I}_{rN} + \tau \boldsymbol{D} \otimes (\overline{\boldsymbol{M}}^{-1}\boldsymbol{A}))^{-1}(\boldsymbol{v} \otimes \boldsymbol{1}_N) = \boldsymbol{v} \otimes \boldsymbol{1}_N$.

Proof. The claim follows from $\boldsymbol{A}\boldsymbol{1} = \boldsymbol{0}$. □

References

[1] M.S. McAfee, O. Annunziata, Cross-diffusion in a colloid–polymer aqueous system, Fluid Phase Equilibria 356 (2013) 46–55. http://dx.doi.org/10.1016/j.fluid.2013.07.014.

[2] A. Vergara, F. Capuano, L. Paduano, R. Sartorio, Lysozyme mutual diffusion in solutions crowded by poly(ethylene glycol), Macromolecules 39 (13) (2006) 4500–4506. http://dx.doi.org/10.1021/ma0605705.

[3] V.K. Vanag, I.R. Epstein, Cross-diffusion and pattern formation in reaction-diffusion systems, Phys. Chem. Chem. Phys. 11 (2009) 897–912. http://dx.doi.org/10.1039/B813825G.

[4] A. Madzvamuse, H.S. Ndakwo, R. Barreira, Cross-diffusion-driven instability for reaction-diffusion systems: analysis and simulations, J. Math. Biol. 70 (4) (2015) 709–743. http://dx.doi.org/10.1007/s00285-014-0779-6.

[5] A. Gerisch, M.A.J. Chaplain, Robust numerical methods for taxis–diffusion–reaction systems: applications to biomedical problems, Math. Comp. Mod. 43 (2006) 49–75. http://dx.doi.org/10.1016/j.mcm.2004.05.016.

[6] D. Becherer, M. Schweizer, et al., Classical solutions to reaction–diffusion systems for hedging problems with interacting Itô and point processes, Annals of Appl. Prob. 15 (2005) 1111–1144. http://dx.doi.org/10.1214/105051604000000846.

[7] K.A. Rahman, R. Sudarsan, H.J. Eberl, A mixed-culture biofilm model with cross-diffusion, Bull. Math. Biol. 77 (2015) 2086–2124. http://dx.doi.org/10.1007/s11538-015-0117-1.

[8] S. Hittmeir, A. Jüngel, Cross diffusion preventing blow-up in the two-dimensional Keller–Segel Model, SIAM J Math. Anal. 43 (2011) 997–1022. http://dx.doi.org/10.1137/100813191.

[9] G. Dziuk, C.M. Elliott, Finite element methods for surface PDEs, Acta Numer. 22 (2013) 289–396. http://dx.doi.org/10.1017/s0962492913000056.

[10] P. Chatzipantelidis, Z. Horváth, V. Thomée, On preservation of positivity in some finite element methods for the heat equation, Comp. Meth. Appl. Math. 15 (2015) 417–437. http://dx.doi.org/10.1515/cmam-2015-0018.

[11] Y-Y Nie, V. Thomée, A lumped mass finite-element method with quadrature for a non-linear parabolic problem, IMA J. Numer. Anal. 5 (1985) 371–396. http://dx.doi.org/10.1093/imanum/5.4.371.

[12] C.M. Elliott, A.M. Stuart, The global dynamics of discrete semilinear parabolic equations, SIAM J Num. Anal. 30 (1993) 1622–1663. http://dx.doi.org/10.1137/0730084.

[13] I. Faragó, J. Karátson, S. Korotov, Discrete maximum principles for nonlinear parabolic PDE systems, IMA J. Numer. Anal. 32 (2012) 1541–1573. http://dx.doi.org/10.1093/imanum/drr050.

[14] X. Li, W. Huang, Maximum principle for the finite element solution of time-dependent anisotropic diffusion problems, Num. Meth. Part. Diff. Eqns. 29 (2013) 1963–1985. http://dx.doi.org/10.1002/num.21784.

[15] C. Lu, W. Huang, J. Qiu, Maximum principle in linear finite element approximations of anisotropic diffusion–convection–reaction problems, Numer. Math. 127 (2014) 515–537. http://dx.doi.org/10.1007/s00211-013-0595-8.

[16] D. Hoff, Stability and convergence of finite difference methods for systems of nonlinear reaction-diffusion equations, SIAM J Num. Anal. 15 (1978) 1161–1177. http://dx.doi.org/10.1137/0715077.

[17] M. Frittelli, A. Madzvamuse, I. Sgura, C. Venkataraman, Lumped finite element method for reaction-diffusion systems on compact surfaces, ArXiv Preprint ArXiv:1609.02741 (2016).

[18] G. Dziuk, Finite elements for the beltrami operator on arbitrary surfaces, Partial Differential Equations and Calculus of Variations 1357 (1988) 142–155. http://dx.doi.org/10.1007/bfb0082865.

[19] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Springer, 2015. http://dx.doi.org/10.1007/978-3-642-61798-0.

[20] E. Hebey, Sobolev Spaces on Riemannian Manifolds, Springer Science & Business Media, 1996. http://dx.doi.org/10.1007/bfb0092907.

[21] M.E. Taylor, Partial Differential Equations. III, Springer-Verlag, New York, 2011. http://dx.doi.org/10.1007/978-1-4419-7049-7.

[22] A.J. Laub, Matrix Analysis for Scientists and Engineers, SIAM, 2005. http://dx.doi.org/10.1137/1.9780898717907.

[23] J. Smoller, Shock Waves and Reaction–Diffusion Equations, Springer Science & Business Media, New York, 1994. http://dx.doi.org/10.1007/978-1-4612-0873-0.

[24] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, Springer-Verlag, Berlin Heidelberg, 1997. http://dx.doi.org/10.1007/978-3-662-03359-3.

[25] M.R. Garvie, C. Trenchea, Finite element approximation of spatially extended predator–prey interactions with the Holling type II functional response, Numer. Math. 107 (2007) 641–667. http://dx.doi.org/10.1007/s00211-007-0106-x.

[26] R.H. Nochetto, C. Verdi, Combined effect of explicit time-stepping and quadrature for curvature driven flows, Numer. Math. 74 (1996) 105–136. http://dx.doi.org/10.1007/s002110050210.

[27] P-A Raviart, The use of numerical integration in finite element methods for solving parabolic equations, Topics Num. Anal. (1973) 233–264.

[28] Q. Du, L. Ju, L. Tian, Finite element approximation of the Cahn–Hilliard equation on surfaces, Comp. Meths. Appl. Mech. Eng. 200 (2011) 2458–2470. http://dx.doi.org/10.1016/j.cma.2011.04.018.

[29] C.M. Elliott, T. Ranner, Evolving Surface Finite Element Method for the Cahn–Hilliard equation, Numer. Math. 129 (2015) 483–534. http://dx.doi.org/10.1007/s00211-014-0644-y.

[30] C. Lubich, D. Mansour, Variational discretization of wave equations on evolving surfaces, Math. Comp. 84 (2015) 513–542. http://dx.doi.org/10.1090/s0025-5718-2014-02882-2.

[31] G. Dziuk, C.M. Elliott, L^2-estimates for the Evolving Surface Finite Element Method, Math. Comp. 82 (2013) 1–24. http://dx.doi.org/10.1090/s0025-5718-2012-02601-9.

[32] O. Lakkis, A. Madzvamuse, C. Venkataraman, Implicit–explicit timestepping with finite element approximation of reaction–diffusion systems on evolving domains, SIAM J Num. Anal. 51 (2013) 2309–2330. http://dx.doi.org/10.1137/120880112.

[33] S. Ghorai, S. Poria, Turing patterns induced by cross-diffusion in a predator-prey system in presence of habitat complexity, Chaos Solitons Fractals 91 (2016) 421–429. http://dx.doi.org/10.1016/j.chaos.2016.07.003.

[34] E. González-Olivares, R. Ramos-Jiliberto, Dynamic consequences of prey refuges in a simple model system: more prey, fewer predators and enhanced stability, Ecological Modelling 166 (2003) 135–146. http://dx.doi.org/10.1016/s0304-3800(03)00131-5.

Please cite this article in press as: M. Frittelli, et al., Lumped finite elements for reaction–cross-diffusion systems on stationary surfaces, Computers and Mathematics with Applications (2017), http://dx.doi.org/10.1016/j.camwa.2017.07.044.